

75108

Delaval, Jan

---

From: Liu, Samuel  
Sent: Friday, September 06, 2002 9:40 AM  
To: Delaval, Jan  
Subject: 09/806382

09/806382 ✓

Hello, Jan,

Please search the following polypeptide sequences for 09/806382:

SEQ ID NOs: 1 and 2.

Have a great weekend!

Very best,

*Samuel Wei Liu*  
Unit 1653  
CM1 9D08  
Mailbox 9B01  
Tel: 306-3483

Jan Delaval  
Reference Librarian  
Biotechnology & Chemical Library  
CM1 1E07 – 703-308-4498  
ian.delaval@uspto.gov

22

---

1

**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delaval  
 Reference Librarian  
 Biotechnology & Chemical Library  
 CM1 1E07 - 703-308-4498  
 jan.delaval@uspto.gov

**STAFF USE ONLY**

	Type of Search	Vendors and cost where applicable
Searcher: <u>Jan</u>	NA Sequence (#) <input checked="" type="checkbox"/>	STN _____
Searcher Phone #: <u>4458</u>	AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>9/19/02</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>9/10/02</u>	Litigation _____	Lexis/Nexis <input checked="" type="checkbox"/>
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <input checked="" type="checkbox"/>
Clerical Prep Time: <u>10</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>15</u>	Other _____	Other (specify) _____





OM of: US-09-806-382a-2 to: PIR\_71.\* out\_format : pirs  
Date: Sep 9, 2002 3:12 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=framer.n2p.model -DEV=xlp  
-O=/cgn2\_1/USPTO/spool/US09806382/runat\_09092002\_143848\_15421/apr-query.fasta.1.752  
-DB=PIR\_71 -FMT=fasta -SUFFIX=pr -GAP=12.000 -GAPEXT=4.000  
-MIMATCH=0.100 -LOOPT=0.000 -LOOPT=0.000 -GAPEXT=4.500  
-GAPEXT=0.050 -XGAP=10.000 -XGAPEXT=0.500 -FGAP=6.000  
-FGAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500 -DEL=6.000  
-DEL=7.000 -START=1 -MATRIX=blomum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pirs -NORM=ext -HEAPSIZE=500  
-MINLEN=0 -MAXLEN=2000000000 -USER=US09806382.@cgn1.1.155  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIME=OUT=120 -WARN\_TIME=OUT=30  
-NO\_XLPRY -WAIT -THREADS=1

## Search information block:

Query: US-09-806-382a-2  
Query Length: 345  
Database: PIR\_71.\*  
Database sequences: 283138  
Database length: 96089334  
Search time (sec): 60.920000

Sequence	Strd	Orig	ZScore	Escore	Len	Documentation
pir1: A42628	+	378.50	689.71	2.1e-30	122	Calgranulin B - bovine (fragment)
pir1: JN0686	+	355.50	648.31	4.5e-28	113	Calgranulin B - rat
pir1: S68242	+	329.50	600.76	2.0e-25	113	Calgranulin B - mouse
pir2: J4712	+	214.50	392.21	1.0e-13	92	S-100 calcium-binding protein A1
pir2: A55406	+	213.50	390.47	1.0e-13	92	Calgranulin C - pig
pir1: BCUHUR	+	177.50	324.55	6.0e-10	92	S-100 protein beta chain [valid]
pir2: A48015	+	176.50	322.61	7.6e-10	91	S-100 protein beta chain - bovin
pir1: S33985	+	176.50	322.44	7.6e-10	92	S-100 protein beta chain - mouse
pir2: A26557	+	173.50	317.23	1.5e-09	92	S-100 protein alpha chain - west
pir2: S24146	+	165.50	302.33	1.0e-08	95	S-100 protein P - human
pir1: BCUHUR	+	165.50	295.10	2.6e-08	94	S-100 protein alpha chain - bovi
pir2: A41988	+	160.50	293.28	3.2e-08	94	S-100 protein alpha chain - huma
pir2: A45135	+	156.50	285.61	8.3e-08	98	S-100 calcium-binding protein A2
pir2: A30129	+	153.50	280.21	1.7e-07	97	S-100 protein - human (fragment)
pir1: J01300	+	153.50	279.78	1.7e-07	102	S-100 protein - lung - bovine
pir2: A48219	+	150.50	274.38	3.4e-07	101	Calgizartin - rabbit
pir1: I37080	+	148.50	272.22	4.3e-07	105	Calvasculin - human
pir2: I46861	+	148.00	279.86	4.4e-07	38	macrophage migration inhibitory
pir2: S06207	+	147.00	271.63	4.8e-07	101	Calvasculin - mouse
pir2: A53217	+	147.00	268.06	7.7e-07	100	placental calcium-binding prote
pir1: I56163	+	146.50	268.14	8.6e-07	89	Calgranulin A - mouse
pir1: JN0685	+	141.50	259.00	2.8e-06	89	Calgranulin A - rat
pir2: S01759	+	140.00	255.17	4.0e-06	101	Calvasculin - rat
pir2: B48219	+	139.00	252.62	5.1e-06	110	S-100 calcium-binding protein A
pir2: A48118	+	139.00	243.90	6.3e-06	306	major epidermal calcium-binding
pir2: J05064	+	138.00	251.77	6.3e-06	98	S-100 calcium-binding protein A1
pir1: S27011	+	137.50	251.59	7.1e-06	90	Calgizartin - rabbit
pir2: B28363	+	135.00	247.01	1.3e-05	90	Calgizartin - rat
pir2: J00330	+	133.00	239.51	1.6e-05	217	26-Kda Ca2+-binding protein, p2
pir2: A43314	+	133.00	243.28	1.6e-05	89	Calgizartin - mouse
pir1: BCUHUR	+	133.00	245.18	1.6e-05	90	Calgizartin - human
pir1: J00663	+	128.00	233.57	6.6e-05	97	calpactin I light chain - chick
pir2: A28489	+	128.00	233.57	6.6e-05	97	calpactin I light chain - mouse
pir2: J05065	+	127.00	231.66	8.4e-05	98	calpactin I light chain - human
pir1: J01139	+	123.00	224.61	0.0002	95	calpactin I light chain - human
pir2: B28489	+	123.00	224.43	0.0002	97	calpactin I light chain - bovine

seq\_name: pir1: A42628  
seq\_documentation\_block:  
calgranulin B - bovine (fragment)  
N: Alternate names: calcium-binding protein MRP-14; macrophage migration inhibitory fa  
in 2  
C: Species: Bos primigenius taurus (cattle)  
C: Date: 30-Sep-1993 #sequence\_revision 23-May-1997 #text\_change 23-May-1997  
C: Accession: B22309; A42628  
R: Tang, T.K.; Hong, T.M.; Lin, C.Y.; Lai, M.L.; Liu, C.H.L.; Lo, H.J.; Wang, M.E.; Ch  
submitted to the Protein Sequence Database, July 1992  
A: Reference number: A22309  
A: Accession: B22309  
A: Status: preliminary  
A: Molecule type: protein  
A: Residues: 1-122 <TRAN>  
R: Dianoux, A.C.; Stasia, M.J.; Garin, J.; Gagnon, J.; Vignais, P.V.  
Biochemistry 31, 5898-5905, 1992  
A: Title: The 23-kilodalton protein, a substrate of protein kinase C, in bovine neutro  
A: Reference number: A42628; M01D:92304974  
A: Accession: A42628  
A: Molecule type: protein  
A: Residues: 4-32, 34-56 <DIA>  
C: Complex: heterodimer and higher complexes with calgranulin A  
C: Superfamily: S-100 protein; calmodulin repeat homology  
C: Keywords: blocked amino end; calcium binding; EF hand; heterodimer; inflammation; p  
F: 50-82/Domain: calmodulin repeat homology <EF1>  
F: 50-82/Domain: calmodulin repeat homology <EF2>

## alignment\_scores:

Quality: 378.50 Length: 109  
Ratio: 4.027 Gaps: 1  
Percent Similarity: 86.239 Percent Identity: 69.725

## alignment\_block:

US-09-806-382a-2 x A42628 ..

Align seg 1/1 to: A42628 from: 1 to: 122

```
13  ATGTGCGAGCTGAGCGACACATGAGACCATCATGACACTTCACCA 62
|||||.....|
1  Metserglnmetglnuserierlleglthrllellasnllephstsl 17
63  ATATCTGTAGAGTGGGCGACCCAGACACCTGAAACAGGGGAATTC 112
|||||.....|
17  nTyrSerValArgLeuGlnHstYrAspThrLeuIlleglnYsGlnSerL 34
113  AAGGCGTGGCGGAAATGATCGCAAAATTTCTCAAGAAAGGATTA 162
|||||.....|
34  YsGlnLeuValGlnYsGlnLeuProHsnHeuLeuYsGlnYsLys 50
163  AATGAAAGGTCAATAGAACACATCATGAGAGCTGAGACAAATGCA 212
|||||.....|
51  AAsGlnAlaIleAlaIleAsnGlnIleMetGlnAspLeuAspThrAsnVal 67
213  CAAGCAAGCTGAGCTGAGAGGATTCATCATGCTGATGCGAGGCTAAC 262
|||||.....|
67  pLysGlnLeuSerPheGlnIleuPheIleMetLeuValAlaArgLeuThr 84
263  GGGCGTCCCGCGAGAGATGACAGAG...GGGAGACAGGCGCTGGCG 309
|||||.....|
84  AlAlaSerHsnIsglnGlnMetHnAsnThrAlaProGlnGlnYsLys 100
310  CASCATTAAGCCAGGCTCGGAGAGGC 336
|||||.....|
101  ArgHnIsglYrProGlnYrGlnYsGln 109
```

seq\_name: p1r1:JN0686

seq\_documentation\_block:

calgranulin B - rat  
N:Alternate names: calcium-binding protein MRP-14; macrophage migration inhibitory factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 03-May-1994 #sequence\_revision 23-May-1997 #text\_change 22-Jun-1999  
A:Accession: JN0686  
R:Manatich, T.; Uchida, I.; Wahl, S.M.; McCartney-Francis, N.  
Biochem. Biophys. Res. Commun. 194, 819-825, 1993  
A:Title: Expression and cloning of migration inhibitory factor-related protein (MRP) 8 and  
A:Reference number: JN0685; MUID:93343942  
A:Accession: JN0686  
A:Molecule type: mRNA  
A:Residues: 1-113 <TMA>  
A:Cross-references: GB:L18948; NID:9488156; PIDN:AAA18214.1; PID:9488157  
C:Genetics:  
A:Gene: MRP14  
C:Complex: heterodimer and higher complexes with calgranulin A  
C:Superfamily: S-100 protein; calmodulin repeat homology  
C:Keywords: acetylated amino end; calcium binding; EF hand; heterodimer; inflammation; F  
F:2-113/Product: calgranulin B #status predicted <MAT>  
F:11-45/Domain: calmodulin repeat homology <EF1>  
F:55-87/Domain: calmodulin repeat homology <EF2>  
F:2/Modified site: acetylated amino end (Ala) (in mature form) #status predicted  
F:80-91/Disulfide bonds: #status predicted

alignment\_scores:

Quality: 355.50 Length: 107  
Ratio: 3.864 Gaps: 1  
Percent Similarity: 85.981 Percent Identity: 64.486

alignment\_block:

US-09-806-382a-2 x JN0686 ..

Align seg 1/1 to: JN0686 from: 1 to: 113

```
16 TCGCAGCTGGAACGACATAGACATCATCAACACCTTCACCAATA 65
|||||
7 SerGlnMetGluArgSerIleSerThrIleIleAsnValPheHisGlnI 23
|||||
66 CTCTGTGAAGCTGGGGCCAGACACCTGGAACCGGGGAATTCAG 115
|||||
23 rSerArgLysTyrGlnHisProAspThrLeuAsnLysAlaGluPheLys 40
|||||
116 AGCTGTGCGAAGATCTGCAAAATTTCTCAAGAGAGATTAAGAT 165
|||||
40 IuMetValAlaAsnLysAspLeuProAsnPheLeuLysArgGluLysArgAsn 56
|||||
166 GAAAGGTGATAGACATCATGAGAGACCTGAGACACAAATGACAGCA 215
|||||
57 GluAsnLeuLeuArgAspIleMetLysAspLeuAspThrAsnGlnAspAs 73
|||||
216 GCAGCTGAGCTTCGAGAGATTCATGCTGATGCGAGAGCTAACCTGG 265
|||||
73 nGlnLeuSerPheGlnLysLysMetMetLeuMetLysLysLeuIlePhe 90
|||||
266 CCTCCACGAGAAATGACAGAGGCTGAGAGGCTTCGGCCACCAACCAT 315
|||||
90 IAcYHisGlnLysLysLeuHisGlnAsnAsnPro...ArgGlnHisAspHis 105
```

seq\_name: p1r1:S68242

seq\_documentation\_block:

calgranulin B - mouse  
N:Alternate names: calcium-binding protein MRP-14; macrophage migration inhibitory factor  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Sep-1996 #sequence\_revision 23-May-1997 #text\_change 22-Jun-1999

C:Accession: S68242; S68272

R:Lagasse, E.; Weissman, I.L.

submitted to the EMBL Data Library, February 1992

A:Description: Mouse MRP8 and MRP14, two intracellular calcium-binding proteins assoc

A:Reference number: S68242

A:Accession: S68242

A:Molecule type: mRNA

A:Residues: 1-113 <LAG>

A:Cross-references: EMBL:M83219; NID:9199807; PIDN:AB07228.1; PID:9199808

R:Raftery, M.J.; Harrison, C.A.; Alewood, P.; Jones, A.; Geczy, C.L.

Biochem. J. 316, 285-293, 1996

A:Title: Isolation of the murine S100 protein MRP14 (14 kDa migration-inhibitory-fact

ding.

A:Reference number: S68272; MUID:96235204

A:Accession: S68272

A:Molecule type: protein

A:Residues: 2-107-95-109 <RAF>

A:Note: 107-His is identified as 3'-methylhistidine: the authors' source for the refe

ylhistidine

C:Genetics:

A:Gene: MRP14

C:Complex: heterodimer and higher complexes with calgranulin A

C:Superfamily: S-100 protein; calmodulin repeat homology

C:Keywords: acetylated amino end; calcium binding; EF hand; heterodimer; inflammation

F:2-113/Product: calgranulin B #status predicted <MAT>

F:11-45/Domain: calmodulin repeat homology <EF1>

F:55-87/Domain: calmodulin repeat homology <EF2>

F:2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental

F:80-91/Disulfide bonds: #status experimental

F:103,105,107/Binding site: zinc (His) #status predicted

F:107/Modified site: 3'-methylhistidine (His) #status experimental

alignment\_scores:

Quality: 329.50 Length: 107  
Ratio: 3.702 Gaps: 1  
Percent Similarity: 83.178 Percent Identity: 59.813

alignment\_block:

US-09-806-382a-2 x S68242 ..

Align seg 1/1 to: S68242 from: 1 to: 113

```
16 TCGCAGCTGGAACGACATAGACATCATCAACACCTTCACCAATA 65
|||||
7 SerGlnMetGluArgSerIleSerThrIleIleAspThrPheHisGlnI 23
|||||
66 CTCTGTGAAGCTGGGGCCAGACACCTGGAACCGGGGAATTCAG 115
|||||
23 rSerArgLysTyrGlnHisProAspThrLeuAsnLysAlaGluPheLys 40
|||||
116 AGCTGTGCGAAGATCTGCAAAATTTCTCAAGAGAGATTAAGAT 165
|||||
40 IuMetValGlnAlaGlnLeuAlaThrPheMetLysLysLysLysArgAsn 56
|||||
166 GAAAGGTGATAGACATCATGAGAGACCTGAGACACAAATGACAGCA 215
|||||
57 GluAlaLeuIleAsnAspIleMetGlnAspLeuAspThrAsnGlnAsp 73
|||||
216 GCAGCTGAGCTTCGAGAGATTCATGCTGATGCGAGAGCTAACCTGG 265
|||||
73 nGlnLeuSerPheGlnLysLysMetMetLeuMetLysLysLeuIlePhe 90
|||||
266 CCTCCACGAGAAATGACAGAGGCTGAC...GAGGCCCTGAGCCACAC 312
|||||
90 IAcYHisGlnLysLysLeuHisGlnAsnAsnProArgLysGlnHisGlnHisSer 106
```

seq\_name: p1r2:J04712

```
seq_documentation_block:
S-100 calcium-binding protein A12 - human
M:Alternate names: calcium-binding amniotic fluid protein 1 (CAAF1); calgranulin C; calg
utrophil protein
C:Species: Homo sapiens (man)
C>Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 13-Aug-1999
C:Accession: Jc4712; Jc4717; Jc4891; S56113; S56114
R:Yamamura, T.; Hicomi, J.; Nagasaki, K.; Suzuki, M.; Takahashi, E.; Saito, S.; Tsukada,
Blochem. Biophys. Res. Commun. 221, 356-360, 1996
A:Title: Human CAAF1 gene - molecular cloning, gene structure, and chromosome mapping.
A:Reference number: Jc4712; MUID:96192053
A:Accession: Jc4712
A:Molecule type: mRNA
A:Residues: 1-92 <YAM>
A:Cross-References: DDBJ:D83657; MID:g1502284; PIDN:BAA12030.1; PID:g1502285
R:Marli, T.; Ertlmann, K.D.; Gallin, M.Y.
Blochem. Biophys. Res. Commun. 221, 454-458, 1996
A:Title: Host-parasite interaction in human onchocerciasis: Identification and sequence
A:Reference number: Jc4717; MUID:96192069
A:Accession: Jc4717
A:Molecule type: protein
A:Residues: 2-92 <MAR>
A:Experimental source: Onchocerca volvulus infecting human tissue
R:Ilig, E.C.; Troxler, H.; Buerjesset, D.M.; Kuster, T.; Markert, M.; Guignard, F.; Hunz
Blochem. Biophys. Res. Commun. 225, 146-150, 1996
A:Title: Amino acid sequence determination of human S100 A12 (P6, Calgranulin C, CGRP, C
A:Reference number: Jc4891; MUID:96332419
A:Accession: Jc4891
A:Molecule type: protein
A:Residues: 2-92 <ILG>
R:Guignard, F.; Manuel, J.; Markert, M.
Blochem. J. 309, 395-401, 1995
A:Title: Identification and characterization of a novel human neutrophil protein related
A:Reference number: S56113; MUID:95351965
A:Accession: S56113
A:Status: preliminary
A:Molecule type: protein
A:Residues: 'X', '4-14', 'X', '16-17', 'XXXX' <GUI1>
A:Experimental source: Isoform 6a
A:Accession: S56114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-21 <GUI2>
A:Experimental source: isoform 6b
C:Comment: This protein is released by activated neutrophils in the course of inflammat
C:Genetics:
A:Gene: GDB:S100A12; P6; MRP6; CGRP; CAAF1
A:Cross-References: GDB:5218374
A:Map position: 1q21-1q21
C:Complex: monomer
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand; monomer; neutrophil; zinc
F:2-92/Product: S-100 calcium-binding protein A12 #status experimental <MAT>
F:6-39/Domain: calmodulin repeat homology <EF1>
F:49-81/Domain: calmodulin repeat homology <EF2>
F:86-90/Region: zinc binding #status predicted

alignment_scores:
Quality: 214.50 Length: 92
Ratio: 2.979 Gaps: 1
Percent Similarity: 78.261 Percent Identity: 46.739

alignment_block:
US-09-806-382a-2 x Jc4712 ..
Align seg 1/1 to: Jc4712 from: 1 to: 92

13 ATGTCCGAGCTGGAGCAACATAGAGACCATCATCAACACCTTCACCA 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MetThrLysLeuGIuGIuHISLeuGIuGIuLLeValAsnIlePheHisGI 17
63 ATACTGTGTGAAGCTGGGGCAGCCAGACACCTTGACACGAGGGGAAATCA 112
```

```
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 nTyIeYValAlaArgLysGIuHISpHeAsPThrLeuSerLysGIuLeuL 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 AAGAGCTGTGGCAAAAGATCTGCAGAAATTTCTCAAGAGGAGAAATAG 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 yGInLeuLeuThrLysGIuLeuAlaAsnThrIle..LysAsnIleLys 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 AATGAAGAAGTCATAGAACACATCATGTGAGGACCTGTGACACAAATGCAGA 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 AsPrySAIValIleAlAspIleAspIlePheGIuLysLeuAspAlaAsnGIuAs 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 CAAGCAGCTGAGCTTCGAGAGACTTCATGCTGATGCGAGGCTAACCT 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 pGIuGIuValAspPheGIuGIuPheIleSerLeuValAlaIleAlaLeuL 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 GGGCTCCAGCAGAGAAGATGCACGAG 288
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 ySAIAlaIleHisThrHisThrLys 91
```

```
seq_name: p1r2:A55406
seq_documentation_block:
calgranulin c - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 23-May-1997
C:Accession: A55406
R:Deil'Angelica, E.C.; Schleicher, C.H.; Santome, J.A.
J. Biol. Chem. 269, 28929-28936, 1994
A:Title: Primary structure and binding properties of calgranulin C, a novel S100-like
A:Reference number: A55406; MUID:95050708
A:Accession: A55406
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-91 <DEL>
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:48-80/Domain: calmodulin repeat homology <EF2>
```

```
alignment_scores:
Quality: 213.50 Length: 91
Ratio: 2.847 Gaps: 1
Percent Similarity: 82.418 Percent Identity: 45.055

alignment_block:
US-09-806-382a-2 x A55406 ..
Align seg 1/1 to: A55406 from: 1 to: 91

16 TCGCAGCTGGAACGCACATAGAGACATCATCAACACCTTCACCAATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 ThrLysLeuGIuAspHisLeuGIuGIuLLeIleAsnIlePheHisGIuTy 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 CTCTGTGAAGCTGGGGCAGCCAGACACCTGTGACACGAGGGGAATTCAAG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 rSerValAlaArgLeuGIuHISrTyAsPThrLeuIleLysArgGIuLeuLys 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 AGCTGTGCGAAAGATCTGCAGAAATTTCTCAAGAGGAGAAATAGAT 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 InLeuIleThrLysGIuLeuPProAsnThrLeu..LysAsnThrLysAsp 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 GAAAGAGTCATPAGAACACATCATGTGAGGACCTGTGACACAAATGCACAA 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 GIuGIuThrIleAspLysIlePheGIuAsnLeuAspAlaAsnGIuAspGI 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 GCAGCTGAGCTTCGAGAGCTTCATCATGCTGATGCGAGGCTAACCTGCG 265
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 uGIuValSerPheLysGIuPheValIleValLeuValThrAspIleuLett 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 CTTCCACAGAGAGATGCACGAG 288
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 hTrAlaHisAspAsnIleHisLys 90
```

```

seq_name: p1r1:BCHUIB

seq_documentation_block:
S-100 protein beta chain
N:Alternate names: neutral S-100 calcium-binding protein beta
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1986 #sequence_revision 06-Jan-1995 #text_change 08-Dec-2000
C:Accession: A38364; A92972; A03076
R:Allore, R.J.; Friend, W.C.; O'Hanlon, D.; Neilson, K.M.; Bauman, R.; Dunn, R.J.; Marks
J. Biol. Chem. 265, 15537-15543, 1990
A:Title: Cloning and expression of the human S100beta gene.
A:Reference number: A38364; MUID:90368757
A:Accession: A38364
A:Molecule type: DNA
A:Residues: 1-92 <ALL>
A:Cross-references: GB:J05600; GB:M59486; NID:9337726; GB:M59487; NID:9337727; GB:M59488
R:Jensen, R.; Marshak, D.R.; Anderson, C.; Lukas, T.J.; Watterson, D.M.
J. Neurochem. 45, 700-705, 1985
A:Title: Characterization of human brain S100 protein fraction: amino acid sequence of S
A:Reference number: A92972; MUID:85291729
A:Accession: A92972
A:Molecule type: protein
A:Residues: 2-92 <GEN>
R:Baudier, J.; Glasser, N.; Haglid, K.; Gerard, D.
Biochim. Biophys. Acta 790, 164-173, 1984
A:Title: Purification, characterization and ion binding properties of human brain S100b
A:Reference number: A90653; MUID:85023393
A:Contents: annotation: metal ion-binding properties
C:Comment: This protein binds p53, tubulin and many other proteins at physiological conc
different affinities exist for both ions on each monomer. Physiological concentrations
ning sites.
C:Comment: The homodimer contains disulfide bonds, but the bond pattern has not been det
C:Genetics:
A:Gene: GDB:S100B
A:Cross-references: GDB:120360; OMTM:176990
A:Map position: 21q22.3-21q22.3
A:Introns: 46/3
A>Note: the first intron occurs before the initiator codon
C:Complex: homodimer; heterodimer with S-100 protein alpha chain (see PIR:BCHUIA)
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: blocked amino end; brain; calcium binding; EF hand; heterodimer; homodimer;
F:2-92/Product: S-100 protein beta chain #status experimental <MAT>
F:6-40/Domains: calmodulin repeat homology <EF1>
F:49-81/Domains: calmodulin repeat homology <EF2>
F:2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #statu
F:19,22,24,27,32/Binding site: calcium (Asp, Glu, Asp, Lys, Glu) #status predicted
F:62,64,66,68,73/Binding site: calcium (Asp, Asp, Glu, Glu) #status predicted

alignment_scores:
Quality: 177.50 Length: 88
Ratio: 2.536 Gaps: 1
Percent Similarity: 79.545 Percent Identity: 39.773

alignment_block:
US-09-806-382A-2 x BCHUIB ..
Align seg 1/1 to: BCHUIB from: 1 to: 92

13 ATGTGCACTGAGCAACAATGAGACCATGATCAACACCTTCACCA 62
|||||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1 MetSerGluLeuIleuLysAlaMetValAlaLeuIleAspValPheHisG1 17
63 ATACTCTGTGAAGCTGGGCAACCAACACCTGAACACGAGGGAATTC 112
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
17 nIyrSerGlyArgGluGlyAspLysHisLysLysSerGluLeuL 34
113 AAGAGCTGGGGAAGAAGATCTCAAAATTTCTCAAGAGAGATTAAG 162
|||||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
34 yscIleuIleAsnGluLeuSerHisPheLeu...GluIleLys 49

```

```

163 AATGAAAGCTCATAGACATCATGAGACCTGGACACAATGACAGA 212
|||||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
50 GluGlnIleValIleAspLysValMetGluThrLeuAspHisG1Thr 66
213 CAAGCAGCTGAGCTTCGAGAGTTTCATCATGTCGATGCGAGCTTAACT 262
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
66 pglIuGlyCysAspPheGlnIlePheMetAlaPheValIleMetValThr 83
263 GGGCTTCCACGAG 276
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
83 hralAcYshISglu 87

seq_name: p1r1:BCBOIB

seq_documentation_block:
S-100 protein beta chain - bovine
N:Alternate names: neurocalcin delta-binding protein S100-beta
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-May-1979 #sequence_revision 14-Nov-1983 #text_change 24-Nov-1999
C:Accession: A91254; B91110; A90075; S54348; A03077
R:Isobe, T.; Okuyama, T.
Eur. J. Biochem. 89, 379-388, 1978
A:Title: The amino-acid sequence of S-100 protein (PAP-I-b protein) and its relation
A:Reference number: A91254; MUID:79045265
A:Accession: A91254
A:Molecule type: protein
A:Residues: 'ESGL', 5-91 <ISO>
A:Experimental source: brain
A>Note: this sequence has since been revised in reference A91110
R:Isobe, T.; Okuyama, T.
Eur. J. Biochem. 116, 79-86, 1981
A:Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a protein
A:Reference number: A91110; MUID:81236562
A:Accession: B91110
A:Molecule type: protein
A:Residues: 1-91 <IS2>
R:Baudier, J.; Gerard, D.
Biochemistry 22, 3360-3369, 1983
A:Title: Ions binding to S100 proteins: structural changes induced by calcium and zin
A:Reference number: A90471; MUID:84000339
A:Contents: annotation: metal ion-binding properties
R:Marshak, D.R.; Umekawa, H.; Watterson, D.M.; Hidak, H.
Arch. Biochem. Biophys. 240, 777-780, 1985
A:Title: Structural characterization of the calcium binding protein S100 from adipose
A:Reference number: A90075; MUID:85278169
A:Accession: A90075
A:Molecule type: protein
A:Residues: 1-91 <MAR>
A:Experimental source: adipose tissue
R:Okazaki, K.; Obata, N.H.; Inoue, S.; Hidak, H.
Biochem. J. 306, 551-555, 1995
A:Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in g
A:Reference number: S54343; MUID:95194333
A:Accession: S54348
A:Molecule type: protein
A:Residues: 56-61, 'N', 63-79, 'V' <OKA>
C:Comment: The S-100 protein is composed of two related polypeptide chains, alpha and
brain proteins, S-100 is also found in a variety of other tissues.
C:Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc
es with different affinities exist for both ions on each monomer. Physiological conce
cium-binding sites.
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: blocked amino end; brain; calcium binding; EF hand; zinc
F:5-39/Domains: calmodulin repeat homology <EF1>
F:48-80/Domains: calmodulin repeat homology <EF2>
F:1/Modified site: blocked amino end (Ser) (probably acetylated) #status experimental
F:18,21,23,26,31/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F:61,63,65,67,72/Binding site: calcium (Asp, Asp, Glu, Glu) #status predicted

alignment_scores:
Quality: 176.50 Length: 87
Ratio: 2.558 Gaps: 1

```

Percent Similarity: 79.310 Percent Identity: 39.080

alignment\_block:

US-09-806-382A-2 x BCB01B ..

Align seg 1/1 to: BCB01B from: 1 to: 91

```
16 TCGCAGCTGGAAAGCAACATAGAGACCATCATCAACACCTTCACCAATA 65
|||||:|||||: : : : : : : : : : : : : : : : : : : : :
1 SerGluLeuGluValAlaValAlaLeuIleAspValPheHisIleu 17
66 CNOTGTGAAGCTGGGCGACCCAGACACCTGAAACCGGGGGAATTCAAG 115
|||||: : : : : : : : : : : : : : : : : : : : :
17 rSerGluYargGluGlyAspLysHisLysLeuLysLysSerGluLeuYsg 34
116 AGCTGTGCGAAAAGATCTGCAAAATTTTCTCAAGAGGAGAAATAGAT 165
|||||: : : : : : : : : : : : : : : : : : : : :
34 lueuIleAsnAsnGluLeuSerHisPheLeu...GluGluIleLysGlu 49
166 GAAAGGCTATGAAACACATCATGAGAGACCTGGACACAAATGACACAA 215
: : : : : : : : : : : : : : : : : : : : :
50 GluGluValAlaAspLysValMetGluThrLeuAspSerAspGlyAspG1 66
216 GCAGCTGAGCTTCGAGAGTTCATCATGCTGATGGCGAGGCTAACCTGGG 265
: : : : : : : : : : : : : : : : : : : : :
66 yGluCysAspPheGluGluGluPheMetAlaPheValAlaMetIleThrTha 83
266 CCTCCACGACGAG 276
||| |||||
83 lAcysHisGlu 86
seq_name: p1r2:A48015
```

seq\_documentation\_block:  
S-100 protein beta chain - mouse  
C:Species: Mus musculus (house mouse)  
C:date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 24-Nov-1999  
C:Accession: A48015  
R:Jiang, H.; Shah, S.; Hilt, D.C.  
J. Biol. Chem. 268, 20502-20511, 1993  
A:title: Organization, sequence, and expression of the murine S100beta gene. Transcript  
A:Reference number: A48015; MUID:93388628  
A:Accession: A48015  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-92 <J1A>  
A:Cross-references: GB:I22144; NID:g404768; PIDN:AAA03075.1; PID:g404769  
C:Genetics:  
A:introns: 46/3  
C:Superfamily: S-100 protein; calmodulin repeat homology  
C:Keywords: acetylated amino end; calcium binding; EF hand  
F:2-92/Product: S-100 protein beta chain #status predicted <MAT>  
F:6-40/Domain: calmodulin repeat homology <EF1>  
F:49-81/Domain: calmodulin repeat homology <EF2>  
F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted  
F:20,22,24,27,32/Binding site: calcium (Gly, Glu, Asp, Lys, Glu) #status predicted  
F:62,64,66,68,70,73/Binding site: calcium (Asp, Asp, Asp, Glu, Asp, Glu) #status predicted

alignment\_scores:

Quality:	176.50	Length:	88
Ratio:	2.558	Gaps:	1
Percent Similarity:	78.409	Percent Identity:	39.773

alignment\_block:

US-09-806-382A-2 x A48015 ..

Align seg 1/1 to: A48015 from: 1 to: 92

```
13 ATGTGCGAGCTGGAAGCAACATAGAGACCATCATCAACACCTTCACCA 62
|||||:|||||: : : : : : : : : : : : : : : : : : : : :
1 MetSerGluLeuGluValAlaMetValAlaLeuIleAspValPheHisG1 17
```

```
63 ATACTGTGTGAAGCTGGGCGACCCAGACACCTTCAGACAGGGGAATTCA 112
|||||: : : : : : : : : : : : : : : : : : : : :
17 ntYrSerGluYargGluGlyAspLysHisLysLeuLysLysSerGluLeu 34
113 AAGAGCTGTGCGAAAAGATCTGCAAAATTTTCTCAAGAGGAGAAATAG 162
|||||: : : : : : : : : : : : : : : : : : : : :
34 ySGLuLeuIleAsnAsnGluLeuSerHisPheLeu...GluGluIleLys 49
163 AATGAAAGGTCATGAAACACATCATGAGAGACCTTGACACCAATGACGA 212
: : : : : : : : : : : : : : : : : : : : :
50 GluGluGluValAlaAspLysValMetGluThrLeuAspGluAspGlyAs 66
213 CAAGAGCTGAGCTTCGAGAGTTCATCATGCTGATGGCGAGGCTAACCT 262
| : : : : : : : : : : : : : : : : : : : :
66 pGlyGluValAspPheGluGluGluPheMetAlaPheValAlaMetValThr 83
263 GGGCCTCCACGAG 276
||| |||||
83 hAlaCysHisGlu 87
seq_name: p1r1:S35985
```

seq\_documentation\_block:  
S-100 protein alpha chain - weatherfish  
C:Species: Misgurnus fossilis (weatherfish)  
C:date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S35985  
R:Ivanenkov, V.V.; Gerke, V.; Minin, A.A.; Plessmann, U.; Weber, K.  
Mol. Cell. Dev. 42, 151-158, 1993  
A:title: Transduction of Ca(2+) signals upon fertilization of eggs; identification of  
A:Reference number: S35985; MUID:94031845  
A:Accession: S35985  
A:Molecule type: protein  
A:Residues: 1-95 <IVA>  
C:Superfamily: S-100 protein; calmodulin repeat homology  
C:Keywords: calcium binding; EF hand  
F:49-81/Domain: calmodulin repeat homology <EF2>

alignment\_scores:

Quality:	176.50	Length:	88
Ratio: <td>2.634<td>Gaps:<td>1</td></td></td>	2.634 <td>Gaps:<td>1</td></td>	Gaps: <td>1</td>	1
Percent Similarity: <td>76.136<td>Percent Identity:<td>43.182</td></td></td>	76.136 <td>Percent Identity:<td>43.182</td></td>	Percent Identity: <td>43.182</td>	43.182

alignment\_block:

US-09-806-382A-2 x S35985 ..

Align seg 1/1 to: S35985 from: 1 to: 95

```
13 ATGTGCGAGCTGGAAGCAACATAGAGACCATCATCAACACCTTCACCA 62
|||||:|||||: : : : : : : : : : : : : : : : : : : : :
1 ValSerGluLeuGluSerAlaMetGluSerLeuIleLysValPheHisTh 17
63 ATACTGTGTGAAGCTGGGCGACCCAGACACCTTCAGACAGGGGAATTCA 112
|||||: ||| ||| ||| : : : : : : : : : : : : : : :
17 rTySerSerLysGluGlyAspLysTyLysLeuSerLysAlaGluLeu 34
113 AAGAGCTGTGCGAAAAGATCTGCAAAATTTTCTCAAGAGGAGAAATAG 162
|||||: : : : : : : : : : : : : : : : : : : : :
34 ySserLeuLeuGluGluGluLeuAsnAspPheLeu...SerAlaSerLys 49
163 AATCAAAAGTCATAGAACACATCATGAGAGACCTTGACACCAATGACGA 212
: : : : : : : : : : : : : : : : : : : : :
50 AspPheMetValAlaGluLysIleMetSerAspLeuAspGluAsnGlnAs 66
213 CAAGAGCTGAGCTTCGAGAGTTCATCATGCTGATGGCGAGGCTAACCT 262
| : : : : : : : : : : : : : : : : : : : :
66 pGlyGluValAspPheGluGluGluPheValAlaLeuValAlaLeuThr 83
263 GGGCCTCCACGAG 276
||| : : : |||
83 aAlaCysAsnGlu 87
```



S-100 protein alpha chain - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 14-Nov-1983 #sequence\_revision 06-Feb-1995 #text\_change 24-Nov-1999  
 C:Accession: A24156; A91110; S54346; A03078  
 R:Kuwano, R.; Maeda, T.; Usui, H.; Araki, K.; Yamakuni, T.; Onshima, Y.; Kurihara, T.; T  
 FEBS Lett. 202, 97-101, 1986  
 A:Title: Molecular cloning of cDNA of S100alpha subunit mRNA.  
 A:Reference number: A24156; MUID:86248083  
 A:Accession: A24156  
 A:Molecule type: mRNA  
 A:Residues: 1-94 <KUM>  
 R:Isobe, T.; Okuyama, T.  
 Eur. J. Biochem. 116, 79-86, 1981  
 A:Title: The amino acid sequence of the alpha subunit in bovine brain S-100 a protein.  
 A:Reference number: A91110; MUID:81236562  
 A:Accession: A91110  
 A:Molecule type: Protein  
 A:Residues: 2-64, 'D', 66-94 <ISO>  
 R:Baudier, J.; Gerard, D.  
 Biochemistry 22, 3360-3369, 1983  
 A:Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc c  
 A:Reference number: A90471; MUID:84000339  
 A:Contents: annotation; metal ion-binding properties  
 R:Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.  
 Biochem. J. 306, 551-555, 1995  
 A:Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia  
 A:Reference number: S54343; MUID:95194333  
 A:Accession: S54346  
 A:Molecule type: Protein  
 A:Residues: 24-33 <OKA>  
 C:Comment: The S-100 protein is composed of two related polypeptide chains, alpha and be  
 brain proteins. S-100 is also found in a variety of other tissues.  
 C:Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc ve  
 es, with different affinities, exist for both ions on each monomer. Physiological concen  
 alcium-binding sites.  
 C:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: blocked amino end; brain; calcium binding; EF hand; zinc  
 F:2-94/Product: S-100 protein alpha chain <Status predicted <MAT>  
 F:7-41/Domain: calmodulin repeat homology <EF1>  
 F:50-82/Domain: calmodulin repeat homology <EF2>  
 F:20,23,25,28,33/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted  
 F:63,65,67,69,74/Binding site: calcium (Asp, Asn, Asp, Glu, Glu) #status predicted

alignment\_scores:  
 Quality: 161.50 Length: 84  
 Ratio: 2.563 Gaps: 1  
 Percent Similarity: 75.000 Percent Identity: 39.286

alignment\_block:  
 US-09-806-382a-2 x BCBOIA ..

Align seg 1/1 to: BCBOIA from: 1 to: 94

```

16 TCGCAGCTGGAAGCGACATGAGACCATCATCAACACCTTCCACCAATA 65
|||||
3 SerGluLeuGluThrAlaMetGluThrLeuIleAsnValPheHisAlaH1 19
66 CTCTGTGAAGCTGGGGCACCAGACACCTGTGACAGGGGAATTCAGA 115
|||||
19 sserGlyLysGluGlyAspLysTyrLysLeuSerLysLysGluLeuLysG 36
116 AGCTGTGCGAAGATCTGCAAAATTTTCTCAAGAAGAGAGATTAAGAT 165
|||||
36 lntLeuLeuGlnThrGluLeuSerGlyPheLeuAspAlaGln...LysAsp 51
166 GAAAGGTCATAGACACATCATGTGAGAGACCTGTGACACAAATGACACA 215
|||||
52 AlaAspAlaValAspLysValMetLysGluLeuAspGluAsnGlyLysAsp 68
216 GCAGCTGAGCTTCGAGAGTTCATCATGCTGATGCGAGGCTAACCCTGG 265
|||||

```

68 ygluValAspPheGlnGluTyrValValLeuValAlaAlaLeuThrValA 85  
 266 CC 267  
 85 la 85

seq\_name: pIR1:BCHUIA

seq documentation\_block:  
 S-100 protein alpha chain - human  
 N:Alternate names: S-100 calcium-binding protein A1 (S100A1)  
 C:Species: Homo sapiens (man)  
 C>Date: 31-Dec-1993 #sequence\_revision 23-May-1997 #text\_change 22-Jun-1999  
 C:Accession: A44470; S16740  
 R:Engelkamp, D.; Schafer, B.W.; Erne, P.; Heizmann, C.W.  
 Biochemistry 31, 10258-10264, 1992  
 A:Title: S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of thr  
 A:Reference number: A44470; MUID:93041710  
 A:Accession: A44470  
 A:Molecule type: mRNA  
 A:Residues: 1-94 <ENG>  
 A:Cross-references: EMBL:X58079; NID:g36175; PIDN:CAA41107.1; PID:g36176  
 A:Experimental source: heart  
 A:Note: sequence extracted from NCBI backbone (NCBI:116494)  
 C:Comment: This protein binds p53, tubulin and many other proteins at physiological c  
 C:Comment: S-100 is an intracellular protein that binds calcium. It binds zinc more t  
 different affinities exist for both ions on each monomer. Physiological concentratio  
 nding sites.  
 C:Comment: Although predominant among the water-soluble brain proteins, S-100 is also  
 C:Genetics:  
 A:Gene: GDB:S100A1; S100A  
 A:Cross-references: GDB:126839; OMIM:176940  
 A:Map position: 1q21-1q21  
 C:Complex: homodimer; heterodimer with S-100 protein beta chain (see PIR:BCHUIB)  
 C:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: brain; calcium binding; EF hand; heterodimer; homodimer; zinc  
 F:7-41/Domain: calmodulin repeat homology <EF1>  
 F:50-82/Domain: calmodulin repeat homology <EF2>

alignment\_scores:  
 Quality: 160.50 Length: 84  
 Ratio: 2.548 Gaps: 1  
 Percent Similarity: 75.000 Percent Identity: 39.286

alignment\_block:  
 US-09-806-382a-2 x BCHUIA ..

Align seg 1/1 to: BCHUIA from: 1 to: 94

```

16 TCGCAGCTGGAAGCGACATGAGACCATCATCAACACCTTCCACCAATA 65
|||||
3 SerGluLeuGluThrAlaMetGluThrLeuIleAsnValPheHisAlaH1 19
66 CTCTGTGAAGCTGGGGCACCAGACACCTGTGACAGGGGAATTCAGA 115
|||||
19 sserGlyLysGluGlyAspLysTyrLysLeuSerLysLysGluLeuLysG 36
116 AGCTGTGCGAAGATCTGCAAAATTTTCTCAAGAAGAGAGATTAAGAT 165
|||||
36 lntLeuLeuGlnThrGluLeuSerGlyPheLeuAspAlaGln...LysAsp 51
166 GAAAGGTCATAGACACATCATGTGAGAGACCTGTGACACAAATGACACA 215
|||||
52 ValAspAlaValAspLysValMetLysGluLeuAspGluAsnGlyLysAsp 68
216 GCAGCTGAGCTTCGAGAGTTCATCATGCTGATGCGAGGCTAACCCTGG 265
|||||
68 ygluValAspPheGlnGluTyrValValLeuValAlaAlaLeuThrValA 85
266 CC 267
85 la 85

```





OM of: US-09-806-382A-2 to: SwissProt\_40:\* out\_format : pfs  
Date: Sep 9, 2002 3:22 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODE=frame-n2p,model -DEV=xlp  
-O=/cgn2\_1/USPTO\_pool/US09806382/runat\_09092002\_143849\_15565/app-query.fasta\_1.752  
-DB=SwissProt\_40 -OPMT=fastan -SUFFIX=exp -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62  
-TRANS=human40.csl -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09806382 @CGNL 1.62 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPHY -WAIT -THREADS=1

## Search information block:

Query: US-09-806-382A-2  
Query length: 345  
Database: SwissProt\_40:\*  
Database sequences: 105224  
Database length: 38719550  
Search time (sec): 32.310000

## score\_list:

Sequence	Strid	Orig	ZScore	EScore	Len	Documentation
SwissProt_40:S109_BOVIN	386.50	630.20	1.7e-27	122	P28783 bos taurus (bovine).	
SwissProt_40:S109_RAT	353.50	577.33	1.6e-24	112	P50116 rattus norvegicus (rat).	
SwissProt_40:S109_MOUSE	329.50	538.39	2.4e-22	112	P31725 mus musculus (mouse).	
SwissProt_40:S109_RABIT	328.00	535.55	3.3e-22	118	P50117 oryctolagus cuniculus	
SwissProt_40:M126_CHICK	216.00	353.77	4.4e-12	119	P28318 gallus gallus (chicken)	
SwissProt_40:S112_PIG	213.50	351.81	7.4e-12	91	P80310 sus scrofa (pig).	
SwissProt_40:S112_HUMAN	209.50	343.32	1.7e-11	91	P80511 homo sapiens (human).	
SwissProt_40:S112_BOVIN	206.50	340.45	3.2e-11	91	P79105 bos taurus (bovine).	
SwissProt_40:S112_RABIT	181.50	300.80	5.8e-09	81	O77791 oryctolagus cuniculus	
SwissProt_40:S108_BOVIN	176.50	291.78	1.6e-08	91	P02638 bos taurus (bovine).	
SwissProt_40:S108_HUMAN	172.50	283.29	3.7e-08	91	P04271 homo sapiens (human).	
SwissProt_40:S108_MOUSE	171.50	283.67	4.6e-08	91	P50114 mus musculus (mouse).	
SwissProt_40:S108_RAT	168.50	278.80	8.6e-08	91	P04631 rattus norvegicus (rat).	
SwissProt_40:S108_HUMAN	165.50	273.60	1.6e-07	95	P25815 homo sapiens (human).	
SwissProt_40:S10A_RAT	162.50	268.90	3.0e-07	93	P35467 rattus norvegicus (rat).	
SwissProt_40:S10A_BOVIN	161.50	267.27	3.7e-07	93	P02639 bos taurus (bovine).	
SwissProt_40:S10A_HUMAN	160.50	265.65	4.5e-07	93	P23287 homo sapiens (human).	
SwissProt_40:S102_HUMAN	154.50	255.59	1.6e-06	97	P29034 homo sapiens (human).	
SwissProt_40:S102_BOVIN	153.50	253.96	2.0e-06	97	P10462 bos taurus (bovine).	
SwissProt_40:S111_RABIT	153.50	253.57	2.0e-06	102	P24480 oryctolagus cuniculus	
SwissProt_40:S10A_MOUSE	152.50	252.67	2.4e-06	93	P56565 mus musculus (mouse).	
SwissProt_40:S104_HUMAN	150.50	248.78	3.6e-06	101	P26447 homo sapiens (human).	
SwissProt_40:S111_HUMAN	149.50	246.85	4.5e-06	105	P31949 homo sapiens (human).	
SwissProt_40:S104_MOUSE	149.00	246.35	5.0e-06	101	P07091 mus musculus (mouse).	
SwissProt_40:S104_BOVIN	147.00	243.18	7.5e-06	100	P35466 bos taurus (bovine).	
SwissProt_40:S108_MOUSE	146.50	243.37	8.4e-06	88	P27005 mus musculus (mouse).	
SwissProt_40:S101_ICPUP	144.50	239.78	1.3e-05	92	P091061 ictalurus punctatus (cat)	
SwissProt_40:S108_RAT	143.50	238.50	1.6e-05	88	P50115 rattus norvegicus (rat).	
SwissProt_40:S106_CHICK	143.50	238.15	1.6e-05	92	P08953 gallus gallus (chicken)	
SwissProt_40:S105_MOUSE	140.50	233.20	2.9e-05	93	P08945 mus musculus (mouse).	
SwissProt_40:S111_MOUSE	140.50	232.79	2.9e-05	98	P50543 mus musculus (mouse).	
SwissProt_40:S104_RAT	140.00	231.75	3.2e-05	101	P05942 rattus norvegicus (rat).	
SwissProt_40:S107_BOVIN	140.00	231.75	3.2e-05	101	P028050 bos taurus (bovine).	
SwissProt_40:S105_HUMAN	139.00	229.45	4.0e-05	110	P33763 homo sapiens (human).	
SwissProt_40:S113_HUMAN	138.00	228.74	4.9e-05	98	P09384 homo sapiens (human).	
SwissProt_40:S106_RABIT	137.50	228.59	5.4e-05	90	P30801 oryctolagus cuniculus	
SwissProt_40:S106_HORSE	137.00	227.61	6.0e-05	92	O77691 equus caballus (horse)	
SwissProt_40:S111_CHICK	135.50	224.44	8.3e-05	101	P24479 gallus gallus (chicken)	
SwissProt_40:S106_RAT	135.00	224.53	9.2e-05	90	P05964 rattus norvegicus (rat).	
SwissProt_40:S106_MOUSE	134.00	223.00	0.0001	89	P14069 mus musculus (mouse).	

seq_name: SwissProt_40:S109_BOVIN	SwissProt_40:S106_HUMAN	134.00	222.91	0.0001	90	P06703 homo sapiens (human)
SwissProt_40:S108_HUMAN	133.50	221.84	0.0001	93 <th>P05109 homo sapiens (human)</th>	P05109 homo sapiens (human)	
SwissProt_40:S110_CHICK	128.00	212.67	0.0004	96 <th>P27003 gallus gallus (chicken)</th>	P27003 gallus gallus (chicken)	
SwissProt_40:S110_MOUSE	128.00	212.67	0.0004	96 <th>P08207 mus musculus (mouse)</th>	P08207 mus musculus (mouse)	

## seq\_documentation\_block:

ID	S109_BOVIN	STANDARD;	PRT;	122 AA.
AC	P28783;			
DT	01-DEC-1992 (Rel. 24, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Calgranulin B (Neutrophil cytosolic 23 kDa protein) (P23) (BEE22)			
DE	(Fragment).			
GN	S100A9.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Esophageal epithelium;			
RX	MEDLINE=93280230; PubMed=8505358;			
RA	Tang T.K., Hong T.-M., Lin C.-Y., Lai M.-L., Liu C.H.L., Lo H.-J.,			
RA	Wang M.-E., Chen L.B., Chen W.-T., Ip W., Lin D.C., Lin J.J.-C.,			
RA	Lin S., Sun T.-T., Wang E., Wang J.L., Wu R., Wu C.-W., Chien S.-i.			
RT	"Nuclear proteins of the bovine esophageal epithelium. I. Monoclonal			
RT	antibody W2 specifically reacts with condensed nuclei of			
RT	differentiated superficial cells."			
RL	J. Cell Sci. 104:237-247(1993).			
RN	[2]			
RP	SEQUENCE OF 4-56.			
RC	TISSUE=Neutrophils;			
RX	MEDLINE=92304974; PubMed=1610833;			
RA	Dianoux A.-C., Stasia M.-J., Garin J., Gagnon J., Vignais P.V.;			
RT	"The 23-kilodalton protein, a substrate of protein kinase C, in			
RT	bovine neutrophil cytosol is a member of the S100 family."			
RT	Biochemistry 31:5898-5905(1992).			
RT	- SUBUNIT: DISULFIDE LINKED HETERODIMER OF A 7/11 kDa AND A 22/23			
CC	kDa SUBUNITS.			
CC	- SUBCELLULAR LOCATION: CYTOPLASMIC; LOOSELY ASSOCIATED TO THE			
CC	CYTOSKELETON.			
CC	- TISSUE SPECIFICITY: FOUND ESSENTIALLY IN PHAGOCYTTIC CELLS.			
CC	- PTM: PHOSPHORYLATED BY PROTEIN KINASE C.			
CC	- MISCELLANEOUS: CONTAINS TWO CA(2+)-BINDING DOMAINS PER PROTEIN.			
CC	- SIMILARITY: BELONGS TO THE S-100 FAMILY.			
CC	- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.			
DR	HSSP: P02638; IMHO.			
DR	InterPro: IPR002048; EF-hand.			
DR	InterPro: IPR001751; S100_CaBP.			
DR	Pfam: PF00036; efhand; 1.			
DR	Pfam: PF01023; S100; 1.			
DR	PROSITE: PS00018; EF_HAND; PARTIAL.			
DR	PROSITE: PS00303; S100_CaBP; 1.			
KW	Calcium-binding; Phosphorylation.			
FT	NON_TER	1		
FT	CA_BIND	19	32	EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT	CA_BIND	63	74	EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT	SEQUENCE	122 AA;	13673 MW;	P3CA8C4806BCCD CRC64;

## alignment\_scores:

Quality:	386.50	Length:	109
Ratio:	4.068 <td>Gaps:</td> <td>1</td>	Gaps:	1
Percent Similarity:	87.156 <td>Percent Identity:</td> <td>70.642</td>	Percent Identity:	70.642

## alignment\_block:

US-09-806-382A-2 x S109_BOVIN	..
Align seg 1/1 to: S109_BOVIN from: 1 to: 122	



```

RX MEDLINE=92223366; PubMed=1373330;
RA Lagasse E., Weissman I.L.;
RT "Mouse MRP8 and MRP14, two intracellular calcium-binding proteins
RL Blood 79:1907-1915(1992).
RN [12]
RC SEQUENCE FROM N.A.
RA NACKEN W.K.F., Lektstrom-Himes J.A., Sorg C., Manitz M.;
RT "Molecular analysis of the mouse S100A9 gene and evidence that the
RT myeloid specific transcription factor C/EBPepsilon is not required for
RT the regulation of the S100A9/A8 gene expression in neutrophils.";
RL J. Cell. Biochem. 80:606-616(2001).
RN [13]
RP SEQUENCE OF 1-9; 75-92 AND 94-108.
RX MEDLINE=96235204; PubMed=8645219;
RA Rattery M.J., Harrison C.A., Alewood P., Jones A., Geczy C.L.;
RT "Isolation of the murine S100 protein MRP14 (14 kDa
RT migration-inhibitory-factor-related protein) from activated spleen
RT cells: characterization of post-translational modifications and zinc
RT binding."
RL Biochem. J. 316:285-293(1996).
CC -1- MASS SPECTROMETRY: MW=12972; MW_ERR=2; METHOD=Electrospray.
CC -1- MISCELLANEOUS: HAS BEEN SHOWN TO BIND CALCIUM.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M83319; AAB07228.1; -.
DR EMBL; AJ250496; CAC14292.1; -.
DR HSSP; P02638; IMHO
MGP; MGI:1338947; S100A9.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001751; S100_CABP.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
DR PROSITE; PS00303; S100_CABP; 1.
KW Calcium-binding; Acetylation; Methylation.
FT INT_MET 0
FT CA_BIND 23 36 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 67 78 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT DISULFD 79 90
FT MOD_RES 1 1 ACETYLATION.
FT MOD_RES 106 106 METHYLATION.
SQ SEQUENCE 112 AA; 12918 MW; ADF810BC175AEC80 CRC64;

```

## alignment\_scores:

Quality: 329.50 Length: 107  
Ratio: 3.702 Gaps: 1  
Percent Similarity: 83.178 Percent Identity: 59.813

## alignment\_block:

US-09-806-382A-2 x S109\_MOUSE ..

Align seg 1/1 to: S109\_MOUSE from: 1 to: 112

```

16 TCGCAGCTGGAAGCAACATAGACATCATCAACACCTTCACCAATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
6 SerGlnMeGluArgSerIleThrThrIleIleAspThrPheHisGlnIy 22
66 CTCTGTGAAGCTGGGGCAGCAGACACCTGGAACAGGGGGGAATTCAAG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
22 rSerArgLysGluGlnGlyHisProAspThrIleuSerLysGluPheArg 39

```

```

116 AGCTGTGCGAAGAAAGATCTGCAAAATTTCTCAGAGAGAGAAATAAAT 165
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
39 lMetValGluAlaGlnIleuAlaThrPheMetLysLysGluLysArgAsn 55
166 GAAAGGTCATAGAACACATCATCATGAGAGACCTGGACCAATATGACGACAA 215
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
56 GluAlaLeuIleAsnAspIleMetGluAspLeuAspLysPheHisGlnAspAs 72
216 CGAGCTGAGCTTCGAGAGATTCATCATGCTGATGGCAGAGCTTAACCTGGG 265
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
72 nGlnIleuSerPheGlnGluGlnCysMetMetLeuMetAlaLysLeuIlePheA 89
266 CTTCCACGAGAGAGATGACAGAGGGTGAC...GAGGGCCCTGGCCACACC 312
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
89 lAcYSHSLysLysLysLysLysLysLysLysLysLysLysLysLysLysSer 105
313 CATAGCAGCAGCCTCGGGGAG 333
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 HisGlyLysGlyCysGlyLys 112
seq_name: SwissProt_40:S109_RABIT

```

## seq\_documentation\_block:

ID S109\_RABIT STANDARD; PRT; 118 AA.

AC P50117;

DT 01-OCT-1996 (Rel. 34, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Calgranulin B (Migration inhibitory factor-related protein 14)

DE (MRP-14) (Fragment).

GN S100A9 OR MRP-14.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI\_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-NEW ZEALAND WHITE; TISSUE-Neutrophils;

RX MEDLINE=9635278; PubMed=8702688;

RA Yang Z., de Veer M.J., Gardiner E.E., Devanish R.J., Handley C.J.,

RA Underwood J.R., Robinson H.C.;

RT "Rabbit polymorphonuclear neutrophils form 35S-labeled S-sulfo-

RT calgranulin C when incubated with inorganic [35S]sulfate.";

RL J. Biol. Chem. 271:19802-19809(1996).

RN [2]

RP SEQUENCE OF 45-82 FROM N.A.

RC STRAIN-NEW ZEALAND WHITE;

RX MEDLINE=94198229; PubMed=8148323;

RA Mori S., Goto K., Goto F., Mutakami K., Ohkawara S., Yoshinaga M.;

RT "Dynamic changes in mRNA expression of neutrophils during the course

RT of acute inflammation in rabbits.";

RL Int. Immunol. 6:149-156(1994).

CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.

CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC EMBL; AF091849; AAC61771.1; -.

DR EMBL; D17404; BAA04227.1; -.

DR HSSP; P02638; IMHO.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR001751; S100\_CABP.

DR Pfam; PF00036; ehand; 1.

DR Pfam; PF01023; S\_100; 1.

DR PROSITE; PS00018; EF\_HAND; 1.

DR PROSITE; PS00303; S100\_CABP; 1.

KW Calcium-binding; Repeat.

```

FT NON_TER 1 1
FT CA_BIND 9 22 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 53 64 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT DOMAIN 103 118 2 X 8 AA TANDEM REPEATS OF G-H-G-H-
FT REPEAT 103 110 1.
FT REPEAT 111 118 2.
SO SEQUENCE 118 AA; 13292 MW; 7496118E21AD5C41 CRC64;

```

```

alignment_scores:
  Quality: 328.00 Length: 102
  Ratio: 3.859 Gaps: 1
Percent Similarity: 83.333 Percent Identity: 63.725

```

alignment\_block:

US-09-806-382a-2 x S109\_RABIT ..

Align seg 1/1 to: S109\_RABIT from: 1 to: 118

```

43 ATGATCAACACTTCCACCAATAGCTGTGAGGTGGGCGACCCAGAC 92
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 llelleasnllephenslgintysetvalatgvalglprorgaspe 17
93 CCGTAACCGAGGGGAATTCAGAGCTGTGCGAAGATCTGCAAAATT 142
||||| ||||||| ||||||| ||||||| ||||||| |||||||
17 rleusergllysgluphelsglleuvalgllylglleuhsasnp 34
143 TTCTCAAGAGGAGATGAAGATGAAAGGTCAAGACATCATGAG 192
||||| ||||||| ||||||| ||||||| ||||||| |||||||
34 heleuylslysglualatrgasplulysalatlleasnaapllemerglu 50
193 GACCTGACACCAATATGCAGACAGACAGCTTGAGAGGATTCATCAT 242
||||| ||||||| ||||||| ||||||| ||||||| |||||||
51 Aspleuasprhtrnsnglnasprlysglnleuserphelglulphelvali 67
243 GGTGATGGCAGGCTTAACCTGGGCTCCACGAGAGATGCAGAGGCT 291
||||| ||||||| ||||||| ||||||| ||||||| |||||||
67 eleuemealalargleuvalhlsalaserhslglumethlsllysasna 84
292 ....GACGAGGGCCCTGGCCACCATGAAGCCAGGCTCGGGGAGGCG 336
||||| ||||||| ||||||| ||||||| ||||||| |||||||
84 larpohlsasprhslsgllyhslserhslglprgllyleuglylser 100
337 ACCCCC 342
|||
101 GlyPro 102

```

seq\_name: SwissProt\_40:M126\_CHICK

seq\_documentation\_block:  
ID M126\_CHICK STANDARD; PRT; 119 AA.

```

AC P28318:
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protealn MRP-126.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=Bone marrow;
RX MEDLINE=92195690; PubMed=1549365;
RA Nakano T., Graf T.;
RT "Identification of genes differentially expressed in two types of
  v-myb-transformed avian myelomonocytic cells.";
RL Oncogene 7:527-534(1992).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN V-MYB-TRANSFORMED MYELOMONOCYTIC
  CELLS.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.

```

```

CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: X61200; -, NOT_ANNOTATED_CDS.
DR HSSP: P04271; IWO.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_CABP.
DR Pfam: PF00036; efhand.1.
DR Pfam: PF01023; S_100.1.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00303; S100_CABP; 1.
KW Calcium-binding.
FT CA_BIND 29 42 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 72 83 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SO SEQUENCE 119 AA; 14065 MW; 2D268DAF6309AD7A CRC64;

```

```

alignment_scores:
  Quality: 216.00 Length: 103
  Ratio: 2.734 Gaps: 2
Percent Similarity: 76.699 Percent Identity: 45.631

```

alignment\_block:

US-09-806-382a-2 x M126\_CHICK ..

Align seg 1/1 to: M126\_CHICK from: 1 to: 119

```

13 ATGTCGAGCTGGAGACCAATAGAACCATATCTCAACACTTCCACCA 62
||||| ||||||| ||||||| ||||||| ||||||| |||||||
11 leuserglleuuglulysalatlleasprvalilleleasprvaliphehslgl 27
63 ATACTGTGAAGCTGGGCGACCCAGACACCTTCAGAGGGGAATTC 112
||||| ||| ||| ||||||| ||||||| |||||||
27 ntyrsetargarglulysprlyasprhtrlethrarglysluleul 44
113 AAGAGCTGTGCGAAAGATCTGCAAAATTTCTCAAGAGAGCAATGAG 162
||||| ||||||| ||||||| ||||||| ||||||| |||||||
44 yslleuulieglulysglleuvalasntylleu...LysHsVallys 59
163 AATGAAGAATCATAGAACACATCATGAGAGCCTGGACCAAAATGAGA 212
||||| ||||||| ||||||| ||||||| ||||||| |||||||
60 Asnglnvalserlleasprglnllephelsasprleuaspnslnlysas 76
213 CAAGCAGCTGAGCTTCGAGAGTTTCATCATGCTGATGCGAGGCTAATCT 262
||||| ||||||| ||| ||| ||||||| ||||||| |||||||
76 pglnglnleuserphelglulvalmetleuileullellyalthrly 93
263 GGGCGCTCCACGAGAGATGCACGAG...GGTACGAGGGCCCTGGCCAC 309
||||| ||||||| ||| ||| ||||||| ||||||| |||||||
93 alalathrhslglnhslleuhsphelcysgluhsprhslnglnhsls 109
310 CACCATAG 318
||||| ||| |||
110 GlnHsGln 112

```

seq\_name: SwissProt\_40:S112\_PIG

seq\_documentation\_block:  
ID S112\_PIG STANDARD; PRT; 91 AA.

```

AC P80310:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcgranulin C (CAGC).
GN S100A12.
OS Sus scrofa (Pig).

```



SEQUENCE FROM N. A. MESTRE-Osorio

16 TCCACCTTGGAAAGCGACATATAGACACATCATATCAACACCTTCCACCATATA 65  
1 ThylaseuGlueuAspHisLeuGluuLysLysLysLysLysLysLysLysLysLysLys 17  
66 CTCTGTGAAGCTGTGGGGACCCACAGACACCTGTACACAGGGGGAATTCAAG 115  
17 rSerValAlaGValAlGlyHisPheAspPThrLeuAsnLysArgGlyLeuLeuysG 34  
116 AGCTGTGCGAAAGACATCTGCACAAATTTTCTTCCAGAGAGGAATTAACAAT 165  
34 LneullehrlhrysglueneuProlysStrhle... GluasnThrLysAsp 49  
166 GAAAGAGTATGAGACACATCATATGAGAGACCTCGACACCAAAATGAGACAA 215  
50 GluPrrThrLleAspLysLlePheGlnAspLeuAspLysAlaAspLysAspG1 66  
216 GCAGCTGACCTTCGAGAGATTCATCATGCTATGAGCGAGGCTAACTCTGG 265

```

66 yAlaValSerPheGluPheValValLeuValSerArgValLeuLysT 83
266 CCTCCACGAGAGATGCAGAG 288
      ::::|::|
      83 hxAIahSIleasPHeIstLys 90
seq_name: SwissProt_40:S112_RABIT

seq_documentation_block:
ID S112_RABIT STANDARD; PRT; 81 AA.
AC 077791;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calgranulin C (CAGC) (Fragment).
GN S100A12.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Neutrophils;
RX MEDLINE=96355278; PubMed=8702688;
RA Yang Z., Devere M.J., Gardiner E.E., Devenish R.J., Handley C.J.,
RA Underwood J.R., Robinson H.C.;
RT "Rabbit polymorphonuclear neutrophils form 35S-labeled S-sulfo-
RT calgranulin C when incubated with inorganic [35S]sulfate."
RL J. Biol. Chem. 271:19802-19809(1996).
CC -1 SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -1 SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: AF091848; AAC61770.1; -.
CC HSSP: P04631; 1B4C.
CC InterPro: IPR002048; EF-hand.
CC InterPro: IPR001751; S100_Casp.
CC Pfam: PF00036; ehand. 1.
CC Pfam: PF01023; S_100; 1.
CC DR PROSITE: PS00018; EF_HAND. 1.
CC DR PROSITE: PS00303; S100_CASP; 1.
CC KW Calcium-binding.
CC FT NON_TER 1 1
CC FT CA_BIND 8 21 EF_HAND 1 (LOW AFFINITY) (BY SIMILARITY).
CC FT CA_BIND 51 62 EF_HAND 2 (HIGH AFFINITY) (BY
CC FT SIMILARITY).
CC SQ SEQUENCE 81 AA; 9401 MW; 95E67A209180CB66 CRC64;

alignment_scores:
Quality: 181.50 Length: 81
Ratio: 2.792 Gaps: 1
Percent Similarity: 80.247 Percent Identity: 41.975

alignment_block:
US-09-806-382a-2 x S112_RABIT ..
Align seg 1/1 to: S112_RABIT from: 1 to: 81
46 ATCAACACCTTCACCACTCTGTGAAGCTGGGGACACACACACCT 95
||||| ||||||| ||||||| ||||||| |||||||
1 IleasnlIlePhehISglntYservalArgtnrgLYhIStYAsPthrc 17
96 GAACACGAGGGAATTCAAAGAGCTGTGGGAAAGATCTGCAAAATTTTC 145
||||| ||||||| ||||||| ||||||| |||||||
17 uselrYsCysGluLeuLYsLYsLeuIlelhrThrgIuLeuValAsnThrl 34

```

```

146 TCAAGAAAGAGATATAGATAAGTCAATAGACACATCATGAGAGC 195
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
34 le...LYsAsnThrlrYsAspGlnAlaThrValAspArgIlePheArgAsp 49
196 CTGCACACAAATGCGACAGACAGCTGAGGCTGCAGGATTCATCATGCT 245
||||| |::|::|::|::|::|::|::|::|::|::|::|::|
50 leuAspGluAspGlyAspPnlSglInValAspPheLYsGluPheLeuSerLe 66
246 GATGCGAGGCTACTGAGCTGCCACGAGAGATGCAGAG 288
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
66 uLeuAlaSerValLeuValThrlAlaHISgluAsnIleHISLYs 80
seq_name: SwissProt_40:S10B_BOVIN

seq_documentation_block:
ID S10B_BOVIN STANDARD; PRT; 91 AA.
AC P02638;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-100 protein, beta chain.
GN S100B.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=79045265; PubMed=710399;
RA Isobe T., Okuyama T.;
RT "The amino-acid sequence of S-100 protein (PAP I-b protein) and its
RT relation to the calcium-binding proteins."
RL Eur. J. Biochem. 89:379-388(1978).
RN [2]
RP REVISIONS TO 1-4.
RX MEDLINE=81236562; PubMed=7250124;
RA Isobe T., Okuyama T.;
RT "The amino-acid sequence of the alpha subunit in bovine brain S-100a
RT protein."
RL Eur. J. Biochem. 116:79-86(1981).
RN [3]
RP SEQUENCE.
RX MEDLINE=85278169; PubMed=4026304;
RA Marshak D.R., Umekawa H., Watterson D.M., Hidaka H.;
RT "Structural characterization of the calcium binding protein s100 from
RT adipose tissue."
RL Arch. Biochem. Biophys. 240:777-780(1985).
RN [4]
RP METAL ION-BINDING PROPERTIES.
RX MEDLINE=84000339; PubMed=6615778;
RA Baudier J., Gerard D.;
RT "Ions binding to S100 proteins: structural changes induced by calcium
RT and zinc on S100a and S100b proteins."
RL Biochemistry 22:3360-3369(1983).
RN [5]
RP CADMIUM-BINDING STUDIES.
RX MEDLINE=91248136; PubMed=2039467;
RA Donato H. Jr., Mani R.S., Kay C.M.;
RT "Spectral studies on the cadmium-ion-binding properties of bovine
RT brain S-100b protein."
RL Biochem. J. 276:13-18(1991).
RN [6]
RP STRUCTURE BY NMR.
RX MEDLINE=96398693; PubMed=8805590;
RA Kilby P.M., van Eldik L.J., Roberts G.C.K.;
RT "The solution structure of the bovine S100B protein dimer in the
RT calcium-free state."
RL Structure 4:1041-1052(1996).
CC -1- FUNCTION: WEARILY BINDS CALCIUM BUT BINDS ZINC VERY TIGHTLY-
CC DISTINCT BINDING SITES WITH DIFFERENT AFFINITIES EXIST FOR BOTH
CC IONS ON EACH MONOMER. PHYSIOLOGICAL CONCENTRATIONS OF POTASSIUM

```





Quality: 172.50 Length: 87  
Ratio: 2.500 Gaps: 1  
Percent Similarity: 79.310 Percent Identity: 39.080

## alignment\_block:

US-09-806-382A-2 x S10B\_HUMAN ..

Align seg 1/1 to: S10B\_HUMAN from: 1 to: 91

```

16 TCGCAGCTGGAACGACATAGACATCATCATACACCTTCCACCAATA 65
   |||||:|||||:  ::  |||||:|||||:|||||:|||||:
1  SerGluLeuGluValAlaMetValAlaLeuLeuAspValPheHisGlnTy 17
66 CTCGTGAAGCTGGGGCACCAGAACCTTGACAGGGGGAATTCAAG 115
   |||||  ::  |||||  |||||:|||||:|||||:|||||:
17 rSerGlyArgGluGlyAspLysHisLysLeuLysLysSerGluLeuLysG 34
116 AGCTGGTGGAAAGATCTGCAAAATTTTCTCAAGAGAGAGATTAAGAT 165
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 LuLeuIleAsnAsnGluLeuSerHisPheLeu..GluGluLeuLysGlu 49
166 GAAAGAGTCATAGAACATCATGAGAGCCTGACACAAATGACAGCAA 215
   ::|||:|||||:  ::|||:|||||  |||||:|||||:|||||:
50 GIngluValAlaLysPlyValMetGluThrLeuAspAsnAspGlyAspG 66
216 GCAGCTGAGCTTCGAGAGCTTCATCATGCTGATGCGCGAGCTAACCTGGG 265
   ::  ::|||:|||||:|||||:  ::|||:|||||  ::|||:
66 yGluCysAspPheGlnGluPheMetAlaPheValAlaMetValThrThra 83
266 CCTCCACGAG 276
   ||  |||||
83 lacYshHisGlu 86

```

seq\_name: SwissProt\_40:S10B\_MOUSE

## seq\_documentation\_block:

```

ID S10B_MOUSE STANDARD; PRT; 91 AA.
AC P50114;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-100 protein, beta chain.
GN S100B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=93388628; PubMed=8376406;
RA Jiang H., Shah S., Hill D.C.;
RT "Organization, sequence, and expression of the murine S100 beta gene.
RT Transcriptional regulation by cell type-specific cis-acting
RT regulatory elements."
RL J. Biol. Chem. 268:20502-20511(1993).
CC -1- FUNCTION: WEAKLY BINDS CALCIUM BUT BINDS ZINC VERY TIGHTLY-
CC DISTINCT BINDING SITES WITH DIFFERENT AFFINITIES EXIST FOR BOTH
CC IONS ON EACH MONOMER. PHYSIOLOGICAL CONCENTRATIONS OF POTASSIUM
CC ION ANTAGONIZE THE BINDING OF BOTH DIVALENT CATIONS, ESPECIALLY
CC AFFECTING HIGH-AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: DIMER OF EITHER TWO ALPHA CHAINS, OR TWO BETA CHAINS, OR
CC ONE ALPHA AND ONE BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: ALTHOUGH PREDOMINANT AMONG THE WATER-SOLUBLE
CC BRAIN PROTEINS, S-100 IS ALSO FOUND IN A VARIETY OF OTHER TISSUES
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DR EMBL: L22144; AA03075.1; -.
DR HSSP: P04631; 1B4C.
DR MGD: MGI:98217; S100b.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_CABP.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF01023; S-100; 1.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00303; S100_CABP; 1.
KW Calcium-binding; Zinc; Metal-binding.
FT INIT MET 0 0
FT CA_BIND 18 31
FT CA_BIND 61 72
FT CA_BIND 91 10597 MW; 2378AAB8BF1C94D CRC64;
SQ SEQUENCE 91 AA; 10597 MW;

```

## alignment\_scores:

Quality: 171.50 Length: 87  
Ratio: 2.522 Gaps: 1  
Percent Similarity: 78.161 Percent Identity: 39.080

## alignment\_block:

US-09-806-382A-2 x S10B\_MOUSE ..

Align seg 1/1 to: S10B\_MOUSE from: 1 to: 91

```

16 TCGCAGCTGGAACGACATAGACATCATCATACACCTTCCACCAATA 65
   |||||:|||||:  ::  |||||:|||||:|||||:|||||:
1  SerGluLeuGluValAlaMetValAlaLeuLeuAspValPheHisGlnTy 17
66 CTCGTGAAGCTGGGGCACCAGAACCTTGAAACAGGGGGAATTCAAG 115
   |||||  ::  |||||  |||||:|||||:|||||:|||||:
17 rSerGlyArgGluGlyAspLysHisLysLeuLysLysSerGluLeuLysG 34
116 AGCTGGTGGCAAAAGATCTGCAAAATTTTCTCAAGAGAGAGATTAAGAT 165
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 LuLeuIleAsnAsnGluLeuSerHisPheLeu..GluGluLeuLysGlu 49
216 GAAAGAGTCATAGAACATCATGAGAGCCTGACACAAATGACAGCAA 215
   ::|||:|||||:  ::|||:|||||  |||||:|||||:|||||:
50 GIngluValAlaLysPlyValMetGluThrLeuAspGluLysPlyAspG 66
216 GCAGCTGAGCTTCGAGAGCTTCATCATGCTGATGCGCGAGCTAACCTGGG 265
   ::  ::|||:|||||:|||||:  ::|||:|||||  ::|||:
66 yGluCysAspPheGlnGluPheMetAlaPheValAlaMetValThrThra 83
266 CCTCCACGAG 276
   ||  |||||
83 lacYshHisGlu 86

```

seq\_name: SwissProt\_40:S10B\_RAT

## seq\_documentation\_block:

```

ID S10B_RAT STANDARD; PRT; 91 AA.
AC P04631;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-100 protein, beta chain.
GN S100B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85037924; PubMed=6093041;
RA Kuwano R., Usui H., Maeda T., Fukui T., Yamanari N., Ohtsuka E.,

```

RA Ikehara M., Takahashi Y. ;  
RT "Molecular cloning and the complete nucleotide sequence of cDNA to  
RT mRNA for S-100 protein of rat brain. ;  
RL Nucleic Acids Res. 12:7455-7465(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kuwano R., Usui H., Maeda T., Araki K., Kurihara T., Yamakuni T.,  
RA Ohtsuka E., Ikehara M., Takahashi Y. ;  
RT "Molecular cloning and nucleotide sequences of cDNA and genomic DNA  
RT for alpha and beta subunits of S100 protein. ;  
RL Taniguchi Symp. Brain Sci. 19:243-255(1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-91359841; PubMed-1653388;  
RA Maeda T., Usui H., Araki K., Kuwano R., Takahashi Y., Suzuki Y. ;  
RT "Structure and expression of rat S-100 beta subunit gene. ;  
RL Brain Res. Mol. Brain Res. 10:193-202(1991).  
RN [4]  
RP SEQUENCE OF 5-91 FROM N.A.  
RX MEDLINE-87137648; PubMed-3818655;  
RA Dunn R., Landry C., O'Hanlon D., Dunn J., Allore R., Brown I.,  
RA Marks A. ;  
RT "Reduction in S100 protein beta subunit mRNA in C6 rat glioma cells  
RT following treatment with anti-microtubular drugs. ;  
RL J. Biol. Chem. 262:3562-3566(1987).  
RN [5]  
RP STRUCTURE BY NMR.  
RX MEDLINE-96387197; PubMed-8794737;  
RA Drohat A.C., Amburgey J.C., Abildgaard F., Starich M.R.,  
RA Baldissari D.M., Weber D.J. ;  
RT "Solution structure of rat apo-S100(beta beta) as determined by NMR  
RT spectroscopy. ;  
RL Biochemistry 35:11577-11588(1996).  
RN [6]  
RP STRUCTURE BY NMR.  
RX MEDLINE-98153156; PubMed-9485423;  
RA Drohat A.C., Baldissari D.M., Rustandi R.R., Weber D.J. ;  
RT "Solution structure of calcium-bound rat S100(beta beta) as  
RT determined by nuclear magnetic resonance spectroscopy. ;  
RL Biochemistry 37:2729-2740(1998).  
RN [7]  
RP STRUCTURE BY NMR.  
RX MEDLINE-99226808; PubMed-10211826;  
RA Drohat A.C., Tjandra N., Baldissari D.M., Weber D.J. ;  
RT "The use of dipolar couplings for determining the solution structure  
RT of rat apo-S100B. ;  
RL Protein Sci. 8:800-809(1999).  
CC -1- FUNCTION: WEAKLY BINDS CALCIUM BUT BINDS ZINC VERY TIGHTLY-  
CC DISTINCT BINDING SITES WITH DIFFERENT AFFINITIES EXIST FOR BOTH  
CC IONS ON EACH MONOMER. PHYSIOLOGICAL CONCENTRATIONS OF POTASSIUM  
CC ION ANTAGONIZE THE BINDING OF BOTH DIVALENT CATIONS, ESPECIALLY  
CC AFFECTING HIGH-AFFINITY CALCIUM-BINDING SITES.  
CC -1- SUBUNIT: DIMER OF EITHER TWO ALPHA CHAINS, OR TWO BETA CHAINS, OR  
CC ONE ALPHA AND ONE BETA CHAIN.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- TISSUE SPECIFICITY: ALTHOUGH PREDOMINANT AMONG THE WATER-SOLUBLE  
CC BRAIN PROTEINS, S-100 IS ALSO FOUND IN A VARIETY OF OTHER TISSUES.  
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X01090; CAA25567.1; -;  
DR EMBL; M54819; AAA42096.1; -;  
DR EMBL; S53527; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; G53522; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; M15705; -; NOT\_ANNOTATED\_CDS.

```

DR PIR; S07357; S07357.
DR PIR; A26557; A26557.
DR PDB; 1SYM; 07-DEC-96.
DR PDB; 1OLK; 11-NOV-98.
DR PDB; 1BAC; 30-DEC-98.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001751; S100_CaBP.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
KW Calcium-binding; zinc; Metal-binding; 3D-structure.
FT INT_MET 0 0
FT CA_BIND 18 31 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY).
FT FT
SO SEQUENCE 91 AA; 10613 MW; 2378AAB8BFF713AD CRC64;

alignment_scores:
    quality: 168.50    length: 87
    ratio: 2.478       gaps: 1
Percent similarity: 78.161    Percent identity: 37.931

alignment_block:
US-09-806-382A-2 x S10B-RAT ..

Align seg 1/1 to: S10B-RAT from: 1 to: 91

16 TCGCAGCTGTGAACGCAACATAGAGACCATCATCACACCTTCCACCATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 SerGluLeuGluGlyAlaMetValAlaLeuIleIleAspValPheHisGlyIu 17
66 CTCTGTGAAGCTGGGGGCAACCCAGACACCCCTGAAACCAAGGGGAATTCAAG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 rSerGlyArgGluGlyAspLysHisLysLeuLysLysSerGluLeuLysG 34
116 AAGCTGTGCGAAAGATCTCGAAATTTTCTCAAGAGAGAAATGAAGAT 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 IuLeuIleAsnAsnGluLeuSerHisPheLeu...GluGluLeuLysGlu 49
166 GAAAGAGTCATAGAACACATCATGAGAGACCTGGACACCAATGACAGACAA 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
50 GInGluValValAspLysValMetGluThrLeuAspGluAspGlyAspGlu 66
216 GCAGCTGAGCTTCGAGAGATTTCATGCTGATGCGAGAGCTAACCTGGG 265
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 yLuoCysAspPheGlnGluPheMetAlaPheValSerMetValThrThra 83
266 CCTCCACAGAG 276
|||:|||||
83 IacYHisGlu 86

seq_name: SwissProt_40::S10B_HUMAN

seq_documentation_block:
ID S10B_HUMAN STANDARD; PRT; 95 AA.
AC P25815;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-100P protein.
GN S100P OR S100E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta; PubMed=1633809;
RX MEDLINE=92339442;
RA Becker T., Gerke V., Kube E., Weber K.;
RT "S100P, a novel Ca(2+)-binding protein from human placenta. cDNA
cloning, recombinant protein expression and Ca2+ binding
```

```
RT properties."
RL Eur. J. Biochem. 207:541-547(1992).
RN [2]
RC SEQUENCE OF 1-91.
RA TISSUE=Placenta;
RX MEDLINE=92171935; PubMed=1540168;
RY Emoto Y., Kobayashi R., Akatsuka H., Hidaka H.;
RT "Purification and characterization of a new member of the S-100
CC protein family from human placenta."
RL Biochem. Biophys. Res. Commun. 182:1246-1253(1992).
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X65614; CAA46566.1; -
DR PIR; S24146; S24146.
DR HSSP; P02638; 1CFP.
DR MIM; 600614; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001751; S100_CaBP.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
DR PROSITE; PS00303; S100_CaBP; 1.
DR KIM; Calcium-binding; Placenta.
FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY).
FT CONFLICT 32 32 E -> T (IN REF. 2).
FT CONFLICT 44 44 F -> E (IN REF. 2).
SQ SEQUENCE 95 AA; 10400 MW; 786E63F3EAC6C1 CRC64;

alignment_scores:
Quality: 165.50 Length: 88
Ratio: 2.364 Gaps: 1
Percent Similarity: 79.545 Percent Identity: 38.636

alignment_block:
US-09-806-382a-2 x S10E_HUMAN ..
Align seg 1/1 to: S10E_HUMAN from: 1 to: 95

13 ATGTGGAGCTGGAACGCACTAGACACATCATCAACACCTTCACCA 62
|||||:|||||: |||:|||||: |||:
1 MetThrGluLeuGluThrAlaMetGlyMetIleLeuAspValPheSer 17
63 ATACTGTGAGCTGGGACACCCAGACACCCCTGAACAGGGGGAATCA 112
|||||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
17 gtyrserglgsergluglglgserThglnThrleuthrlysglgluLeuL 34
113 AAGAGCTGTGGGAAAGATCTGCAAAATTTCTCAAGAGGAGAAATAG 162
|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
34 yvalleuMetglulysgluleuProglyPheLeu...GlnserLyllys 49
163 AATGAAAAGCTATGAAACATCTGAGAGAGCTGACACACAAATGAGA 212
|||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
50 AspLysAspAlaValAspLysLeuLeuLysAspLeuAlaAsnGlyAs 66
213 CAGAGAGCTGAGCTGAGAGAGTTCATCATGCTGAGAGGAGGCTAACT 262
| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
66 palaglnValAspPheSerGluPheIleValPheValAlaIleIleTrn 83
263 GGGCTCCACAGAG 276
||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
83 eraLacysHlsllys 87
```

```
seq_name: SwissProt_40::S10A_RAT
seq_documentation_block:
ID S10A_RAT STANDARD; PRT; 93 AA.
AC P35457;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-100 protein, alpha chain.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Song W.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 10-93 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=92076235; PubMed=1742602;
RA Zimmer D.B., Song W., Zimmer W.E.;
RT "Isolation of a rat S100 alpha cDNA and distribution of its mRNA in
RL rat tissues."
RL Brain Res. Bull. 27:157-162(1991).
CC -1- FUNCTION: WEAKLY BINDS CALCIUM BUT BINDS ZINC VERY TIGHTLY.
CC DISTINCT BINDING SITES WITH DIFFERENT AFFINITIES EXIST FOR BOTH
CC IONS ON EACH MONOMER. PHYSIOLOGICAL CONCENTRATIONS OF POTASSIUM
CC ION ANTAGONIZE THE BINDING OF BOTH DIVALENT CATIONS, ESPECIALLY
CC AFFECTING HIGH-AFFINITY CALCIUM-BINDING SITES.
CC -1- SUBUNIT: DIMER OF EITHER TWO ALPHA CHAINS, OR TWO BETA CHAINS, OR
CC ONE ALPHA AND ONE BETA CHAIN.
CC -1- TISSUE SPECIFICITY: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: ALTHOUGH PREDOMINANT AMONG THE WATER-SOLUBLE
CC BRAIN PROTEINS, S-100 IS ALSO FOUND IN A VARIETY OF OTHER TISSUES.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U26358; AAB53657.1; -.
DR EMBL; U26357; AAB53657.1; JOINED.
DR EMBL; S68809; AAB20539.2; -.
DR HSSP; P04631; 1B4C.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001751; S100_CaBP.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
DR KIM; Calcium-binding; zinc; Metal-binding.
FT INIT_MET 0 0 BY SIMILARITY.
FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY).
FT CONFLICT 13 13 N -> H (IN REF. 2).
FT CONFLICT 56 56 K -> R (IN REF. 2).
SQ SEQUENCE 93 AA; 10429 MW; D2959A95EC0651A9 CRC64;

alignment_scores:
Quality: 162.50 Length: 84
Ratio: 2.579 Gaps: 1
Percent Similarity: 75.000 Percent Identity: 40.476
```



OM of: US-09-806-382a-2 to: SPTREMBL\_19:\* out\_format : pfs

Date: Sep 9, 2002 3:24 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

#### Command line parameters:

```
-MODEL=frimaf_n2p.model -DEV=xlp  
-O=/cgml1/USPRO.spool/US09806382/runtat_09092002.143849_15535/app-query.fasta.1.752  
-DB=SPTREMBL_19 -OFMT=fastin -SUFFIX=rspt -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -ICOPC=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsun62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct  
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NOM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09806382.ecgml1_279 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV=TIMEOUT=120 -MAIN_TIMEOUT=30 -NO_XLPEXY -WAIT -THREADS=1
```

#### Search information block:

```
Query: US-09-806-382a-2  
Query length: 345  
Database: SPTREMBL_19:*  
Database sequences: 362222  
Database length: 172994929  
Search time (sec): 104.890000
```

#### Score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
SP-rodent:Q90VR5	+	256.50	479.27	2.9e-18	83   Q90VR5 rattus sp. calprotectin
SP-vertebrate:Q93395	+	179.50	333.95	2.9e-10	101   Q93395 salvelinus fontinalis (b
SP-rodent:Q92573	+	172.50	321.70	1.6e-09	92   Q92573 cricetus griseus (chib
SP-rodent:Q91V77	+	156.50	291.67	7.1e-08	94   Q91V77 m.11 days embryo cdna, r
SP-human:Q90R83	+	155.50	289.45	9.1e-08	98   Q90R83 homo sapiens (human). s
SP-mammal:Q9TR16	+	154.50	290.50	1.1e-07	70   Q9TR16 bos taurus (bovine). cor
SP-human:Q94401	+	154.50	280.86	1.3e-07	213   Q94401 homo sapiens (human). d
SP-human:Q01720	+	154.00	271.08	1.6e-07	591   Q01720 homo sapiens (human). f
SP-human:Q94402	+	154.00	269.78	1.6e-07	687   Q94402 homo sapiens (human). d
SP-rodent:Q9D3M4	+	152.50	284.21	1.9e-07	94   Q9D3M4 mus musculus (mouse). s
SP-mammal:Q9TV56	+	148.00	275.20	5.5e-07	101   Q9TV56 canis familiaris (dog).
SP-rodent:Q9U108	+	142.50	267.07	2.0e-06	79   Q9U108 mus musculus (mouse). s
SP-human:Q03331	+	141.00	240.57	3.0e-06	1218   Q03331 homo sapiens (human). f
SP-vertebrate:Q9PSF6	+	140.50	262.02	3.3e-06	92   Q9PSF6 ictalurus punctatus (chat
SP-rodent:Q9R2B7	+	135.00	252.05	1.2e-05	89   Q9R2B7 rattus norvegicus (rat).
SP-vertebrate:Q9YH57	+	135.00	244.32	1.3e-05	217   Q9YH57 rana catesbeiana (bull f
SP-human:Q9ICJ1	+	133.00	257.74	1.7e-05	30   Q9ICJ1 homo sapiens (human). cal
SP-human:Q90B63	+	118.00	203.47	0.0009	495   Q90B63 homo sapiens (human). th
SP-human:Q96F06	+	117.00	217.21	0.0009	103   Q96F06 homo sapiens (human). s
SP-rodent:Q91X65	+	111.50	209.14	0.0033	80   Q91X65 mus musculus (mouse). s
SP-human:Q941E2	+	106.00	197.96	0.0125	89   Q941E2 homo sapiens (human). psc
SP-human:Q9HCY8	+	103.50	191.94	0.0232	104   Q9HCY8 homo sapiens (human). s
SP-vertebrate:Q9ST59	+	100.50	170.44	0.0582	652   Q9ST59 drosophila melanogaste
SP-rodent:Q9VSH4	+	95.50	169.80	0.0587	702   Q9VSH4 drosophila melanogaste
SP-rodent:Q9D208	+	95.50	177.02	0.1570	104   Q9D208 mus musculus (mouse). 11
SP-rodent:Q9D1D8	+	95.50	174.57	0.1570	104   Q9D1D8 mus musculus (mouse). 11
SP-rodent:Q9D708	+	95.00	157.02	0.2261	940   Q9D708 mus musculus (mouse). 23
SP-human:Q00405	+	92.00	156.30	0.4352	535   Q00405 arabis thaliana (mc
SP-plant:Q38873	+	92.00	151.42	0.4634	940   Q38873 arabis thaliana (mc
SP-human:Q96Q00	+	91.50	153.85	0.5003	638   Q96Q00 homo sapiens (human). th
SP-rodent:Q87114	+	91.00	171.70	0.4430	73   Q87114 oryctolagus cuniculus (ra
SP-vertebrate:Q18887	+	89.50	156.09	0.7475	320   Q18887 caenorhabditis elegans
SP-virus:Q9PWX1	+	89.00	147.68	0.9275	759   Q9PWX1 human herpesvirus 6p. dt
SP-organella:Q21351	+	88.50	154.53	0.9459	309   Q21351 entadira herklotzi. nadh
SP-plant:Q93759	+	88.50	149.86	1.00	530   Q93759 arabidopsis thaliana (mc
SP-plant:Q42438	+	88.50	149.81	1.01	533   Q42438 arabidopsis thaliana (mc
SP-virus:Q9W9G4	+	88.00	147.20	1.16	647   Q9W9G4 human herpesvirus 6. dt2
SP-bacteria:Q9REK6	+	87.50	158.81	1.11	152   Q9REK6 micrococcus luteus (micr
SP-mammal:Q9T5B1	+	87.00	166.69	1.12	55   Q9T5B1 bos taurus (bovine). calv

SP-rodent:Q91WZ7	+	84.50	149.74	2.38	227   Q91WZ7 rattus norvegicus (ra
SP-plant:Q23184	+	83.50	150.54	2.92	167   Q23184 arabidopsis thaliana
SP-human:Q96BA4	+	83.00	149.15	3.31	176   Q96BA4 homo sapiens (human).
SP-human:Q95616	+	83.00	145.51	3.47	268   Q95616 homo sapiens (human).

seq\_name: SP-rodent:Q90VR5

```
seq_documentation_block:  
ID Q90VR5 PRELIMINARY; PRT; 83 AA.  
AC Q90VR5;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE CALPROTECTIN LARGER COMPONENT MRP-14 (FRAGMENT).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI_TaxID=10118;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9539365; PubMed=7665986;  
RA Yui S., Mikami M., Yamazaki M.;  
RT "Purification and characterization of the cytotoxic factor in rat  
peritoneal exudate cells: its identification as the calcium binding  
protein complex, calprotectin."  
RL J. Leukoc. Biol. 58:307-316(1995).  
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
DR HSSP: P02638; IMHO.  
DR InterPro: IPR002048; EF-hand.  
DR InterPro: IPR001751; S100_CABP.  
DR Pfam: PF00036; efhand; 1.  
DR PROSITE: PS00303; S100_CABP; 1.  
KW Calcium-binding.  
SQ SEQUENCE 83 AA; 9811 MW; 2E1204E5DD72C418 CRC64;
```

#### alignment\_scores:

Quality:	256.50	Length:	83
Ratio:	3.664	Gaps:	1
Percent Similarity:	84.337	Percent Identity:	60.241

#### alignment\_block:

US-09-806-382a-2 x Q90VR5 ..

Align seg 1/1 to: Q90VR5 from: 1 to: 83

```
88 GACACCCGGAACAGGGGGAATTCAGAGAGCTGGGGAAGATTCGCA 137  
|||||  
1 ASPHTLDSALNLSALAGLURHEYSGLIMETVALSNLSASRLPR 17  
138 AATTTTCTCAAGAGAGAGATTAAGTAAGTAAGTAAGTACATCA 187  
|||||  
17 GANRHEULYSATGULYSATGANSGLNLSLNULEULATGSLPEM 34  
188 TGAGAGACCTGACACAAATGACACAAAGACGCTTGAGAGACTTC 237  
|||||  
34 ETGLASPLDSALSPRLHANSGLNLSASRLPRHEGLNGLCYCS 50  
228 ATCATGATGATGGAGAGGCTAATCGGGCTCCGACGAGAAAGATGCA 287  
|||||  
51 MWTGASGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 336  
288 GAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 336  
|||||  
67 WANSNLSRPR...ARGGLNLSASRLHNSARGNLSGLLYSGLYUGSLY 82  
seq_name: SP-vertebrate:Q93395  
seq_documentation_block:  
ID Q93395 PRELIMINARY; PRT; 101 AA.  
AC Q93395;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)
```

```

JT 01-DEC 2001 (Tremblrel. 19, last annotation update)
DE S100B.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Cranista; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuge O., Yamakawa Y., Nishijima M.;
RT "enhancement of transport-dependent decarboxylation of
phosphatidylserine by S100B protein in permeabilized Chinese hamster
ovary cells." ;
RL J. Biol. Chem. 0:0-0(2001) .
DR EMBL; AB056121; BAB43945.1 ;
SQ SEQUENCE   92 AA;  10749 MW;  AF50107EC2BEDF68 CNC64;

alignment_scores:
      Quality: 172.50          Length: 88
      Ratio: 2.537              Gaps: 1
Percent Similarity: 77.273    Percent Identity: 38.636

alignment_block:
US-09-806-382A-2 x Q925T3 ..

Align seg 1/1 to: Q925T3 from: 1 to: 92

13 ATGTGCGAGCTGGAAACGAACATAGAGACCATTCAATCAACACCTTCACCA 62
|||||.....::: :::::::::::::: |||||
```

```

alignment_scores:      Quality: 172.50      Length: 88
                       Ratio: 2.537      Gaps: 1
Percent Similarity: 77.273      Percent Identity: 38.636

Alignment_block:
US-09-806-382A-2 x Q925T3      ..

Align seg 1/1 to: Q925T3      from: 1      to: 92

13 ATGTGCAGCTGGAGCGCAACATAGACATCATCAACACCTTCCACCA 62
   |||||:|||||:::  ::  :::::||||:  |||||
1 MetsergltuleGulblyslametalalaleuIleapIleaphenISGI 17

63 ATATCTGTGAAGTGGGGGCAACCCAGACACCTTGAAACAGGGGGATTTCA 11
   |||||  ::  |||  |||:|||||:|||||:
17 nuyisergluayrglugluayasplyshIslyslendylsysserltuleul 34

```

```

alignment_block:
US-09-806-382A-2 x Q925T3 ..
Align seg 1/1   to: Q925T3 from: 1   to: 92

13 AATGTGGCAGCTGGAAACGCAACAATAAGACATCATCAACACTTTCCACCA 62
|||||:::||||||::: :: ::::|||||::: ||||||
1 MetSerGIuleLengIubLysLaMetValAlaLeuIIeaSPILeePhnIsgI 17
63 ATACCTGTGGAAGCTGGGGGCACCAGAACCCCTGAACAGGAGGGAATTCA 11
|||||::: ::: ||| |||:::|||||::: |||
17 nTyteerGIeulaugIucIGluYaarPlyrNisLySyleuLySySeerGIuleul 34

113 AAGACCTGGTGCGAAAGAATCTGCAAAAATTTTTCTCAGAAGAGCAATAG 16
|||||:::|||||:::|||||:::|||||::: |||
34 ySGIleuIIeuLeksnsngIueusertNisPhleu...GluguIIeuls 49
```

```

Align seg 1/1 to: Q925T3 from: 1 to: 92

13 ATGTGCGACGTGGAAACGCAACATAGAGACCATCATCAACACCTTCCACCA 62
   |||||:::|||||::: :: :::::|||||::: |||||
1 MetserGluLeuGlnLysAlaMetValAlaLeuIleAspIlePheNIsgI 17
63 ATACTCTGTGAAGCTGGGGGCGACCCGACACACCTGAACACGAGGGGGAATGCA 11
   |||||::: ||| |||||:::|||||:::
17 nTyzSerGluArgGlnGluAspLysnIshLysLeuLysLysSerGluLeuL 34
113 AAGCACTGGTGGGAAAGATCTGCAAAATTTTCTCAAGACGAGAAATAG 16
   |||||:::|||||:::|||||::::::||||
34 ysgIleuIleLysnAsnGlnLeuSerHisPheLeu...GluGluIleLys 49
163 AATGAAAGGTCATATGAACACATCTATCGAGGACCTGGACACACAATGCGA 21

```

```

13  ATGTGGACGCTTGAAACGGCAACATCATGAGCAATCATCAACACCTCTACCA 62
14  |||||:::|||||::: |||||:::|||||::: |||||:::|||||::: |||||:::
15  1 MetSerGluLeuGluIuylsAlaMetValAlaLeuIleAspIlePheHisGI 17
16  63 ATACTCTGTGAAGCTGGGGCCACCCAGACACCTGAAACCGAGGGGGATTCA 11
17  |||||::: |||||::: |||||:::|||||::: |||||:::|||||:::
18  17 nTySerGluGlyArgGluGlyAspLysHisIlysLeuIylsSerGluLeu 34
19  113 AAGCACTGGTGGCAAAAGATCTGCAAAATTTTCTCAAGAGAGAAATAG 16
20  |||||:::|||||:::|||||:::|||||::: |||||::: |||||:::
21  34 ySGluLeuIleAsnGlnIleuSerHisPheLeu...GluIuIleIys 49
22  163 AATGAAAAGTCAATGAAACATCATGTGAGGACCTGGAACACAAATGAGA 21
23  |||||:::|||||::: |||||::: |||||::: |||||::: |||||:::
24  50 GluGlnGluValValAspLysValMetCysIuThrIleuAspGluAspGlyAs 66
25  213 CAAGAGCGTGAACCTTCGAGAAATTCATCATGGTATGTGGCGAGAAACCT 26
26  |||||:::|||||::: |||||::: |||||::: |||||::: |||||:::
27  66 polyIuIuYsAspPheGlnGluPheMetAlaPheValIserMetValThrT 83
28  263 GGGCTCCACAGAG 276

```

```

63 ATACGTCTGTGAAGCTGGGGCACCAGACACCCTGAACCCAGGGGGATTTCA 11
||||| ... ||| |||:.....:|||||
17 nTytterGtLyaAgGluGluYAspLysHisLysLeuYsLysSerGluLeu 34
|||||:.....:|||||:.....:|||||
113 AAGAGCTGGTGGCAAAAGATCTGCAAAATTTTCTCAAGAGAGAGATAAG 16
|||||:.....:|||||:.....:|||||
34 ySgluLeuIleAsnSncgluLeuSerHisPheLeu...GluGluLeuLys 49
|||||:.....:|||||:.....:|||||
169 ATGTAAAGCTCATATGAACACATCTATGGAGGACCTGGACACACAATGCGA 21
|||||:.....:|||||:.....:|||||
50 GluGluGluValValAspLysValMetGluThrLeuAspGluAspGluAs 66
|||||:.....:|||||:.....:|||||
213 CAAGACGCTGAAGCTTCGAGAGATTCATCATGCTGATGCGAGAGCTAACCT 26
|:::|||||:.....:|||||:.....:|||||
66 polyGluLysAspRheGluGluPheMetAlaPheValAspMetValThr 83
|||||:.....:|||||:.....:|||||
263 GGGCTCCCCAGCAG 276
83 hTAlaLysHisGlu 87
seq_name: sp_r0dent.09177

```

```

113 AAGAGCTGGTGGCAAAAGATCTGCACAAATTTTCTCACAAGAGAGATAAG 16
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
34 ysgluleuilelksnmsngluuenseerthlspheleu...glugluilelys 49
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
163 AATGAAAGAGTCATATGAACACATCATCTGGAGGACTGGACACACAATGCAGA 21
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
50 gluglmgluvalvalaspysvalmetcglutlrhluaspsgluaspsglyas 66
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
213 CAAGAGGCTGAGACTTCGAGGAGTTCATGCTGATGCTGAGGAGCTAAACT 26
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
66 pelyluglucysasprheglngluhmetalaiphevalisermetvalthrt 83
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
263 GGGCTCCACGAG 276
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
83 hralacysrhlsclu 87
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
seq_name: sp._rodent:091V77

seq_documentation_block:
AC Q91V77 PRELIMINARY; PRT; 94 AA.
Q91V77;
01-DEC-2001 (TrEMBLrel. 19, Created)

```

```

163 AATTGAAAGGTCATGAGACACATCATGTGGAGGACCTGGACACAAATGACGAA 21
      :::::::::::::::::::: ::||| ||||| ::::::|
50 GUGGAGGUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 66
      :::::::::::::::::::: ::||| ||||| ::::::|
213 CAGGACGCTGACGCTTCGAGAGTTCATGATGCTATGCTGGCGAGGCTAACCT 26
      | :: :::::::::::::::::::: :::::::::::: ::|||
66 polylucyasaapheglingluPhemeLaIaPhevaIserMetVaIthrt 83
      ::||| |||||
263 GGGCTCCCGACGAG 276
      ||| |||||
83 hfaIcysHnIsglu 87

seq_name: sp_r0dent01:Q91V77

seq_documentation_block:
ID Q91V77 PRELIMINARY; PRT; 94 AA.
AC Q91V77;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, last annotation update)
DE 11 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY.
DE CLONE:270088BD09, FULL INSERT SEQUENCE (5100 CALCIUM BIND

```

```

50  GUGGNGGUGUValValAlaSprValValMetCylurThrlLeuAspGluAspCylas 66
213  CAGAGCAAGCTGAGCTTCGAGAGAGTTCATATGCTGTATGGCAGAGCTAACCT 26
    |   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
66  polyGluCysAspPheGlnGlnPheMetAlaPheValAspSerMetValAlaThrT  83
263  GGGCTGCTCCACAGAG 276
    ||  |||||
83  hFALCysHISglu 87

seq_name: sp_r0dent01:Q91V77

seq_documentation_block:
ID      Q91V77      PRELIMINARY;      PRT;      94  AA.
AC      Q91V77;
DT      01-DEC-2001 (TEMBLRel. 19, Created)
DT      01-DEC-2001 (TEMBLRel. 19, Last sequence update)
DT      01-DEC-2001 (TEMBLRel. 19, Last annotation update)
DE      11 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE      ALONE.2700088BD09, FULL INSERT SEQUENCE (S100 CALCIUM BIND
DE      AL1) (ADULT MALE KIDNEY CDNA, RIKEN FULL-LENGTH ENRICHED L
DE      CLONE.0610031F03, FULL INSERT SEQUENCE).
DE      S100A1.

```

```

213 CAAAGGACGTGGACTTCGAGGAGTTATCGTGTATGGCAGGCGTAACCT 20
| : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: |||
66 polydUcysAapRhpneGlnGuIuphemeAlaPheValIserMetValThrT 83
263 GGCCTCCACGACG 276
||| |||||
83 hrlAcysHISglu 87
seq_name= sr_r0dent:c91v77
seq_documentation block:
ID_Q91V77 PRELIMINARY; PRT; 94 AA.
AC Q91V77;
DT 01-DEC-2001 (TREMBLrefl_19, Created)
DT 01-DEC-2001 (TREMBLrefl_19, last sequence update)
DT 01-DEC-2001 (TREMBLrefl_19, last annotation update)
DE 11 DAYS EMBROID CNLA, RIKEN FULL-LENGTH ENRICHD LIBRARY,
DE CLONE:z270008BD09, FULL INSERT SEQUENCE (S100 CALCIUM BIND
DE A1) (ADULT MALE KIDNEY CDNA, RIKEN FULL-LENGTH ENRICHD L
DE CLONE:0610031F03, FULL INSERT SEQUENCE).
GN S100AL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Eutele
```

```

263 GGGCTCCACGAG 276
      ||| |||||
      83 hralacynhisglu 87

seq_name: sp_rident:Q91V77

seq_documentation_block:
ID:Q91V77      PRELIMINARY;      PRT;      94 AA.
AC Q91V77;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE 11 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2700088D09, FULL INSERT SEQUENCE (5100 CALCIUM BIND
DE A1) (ADULT MALE KIDNEY CDNA, RIKEN FULL-LENGTH ENRICHED L
DE CLONE:0610031F03, FULL INSERT SEQUENCE).
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
OC Mammalia; Eulhiera; Rodentia; Sciurognathi; Muridae; Muril
OC NBI_TaxID=10090;

```

```

seq_name: sp.trodent:091V77
seq_documentation_block:
ID 091V77          PRELIMINARY;          PRT;          94 AA.
AC 091V77;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 11 DAYS EMBRIO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2700088D09, FULL INSERT SEQUENCE (S100 CALCIUM BIND
DE A1) (ADULT MALE KIDNEY CDNA, RIKEN FULL-LENGTH ENRICHED L
DE CLONE:0610031F03, FULL INSERT SEQUENCE).
GN S100a1.
OS Mus musculus (Mouse).
OC Eukaryotes, Metazoa, Chordata, Craniata, Vertebrata, Eutele
OC Mammalia, Eutheria, Rodentia, Sciurognathii, Muridae, Murit
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.

```

```
seq_name: sp1000000091V77
seq_documentation_block:
  ID: 091V77
  PRELIMINARY:
  PRT:
  94 AA.
AC: 091V77;
DT: 01-DEC-2001 (TEMBLrel. 19, Created)
DT: 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT: 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE: 11 DAEs EMBRIO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE: CLONE:z2700088D09, KUDY INSERT SEQUENCE (5100 CALCIUM BIND
DE: A) (ADULT MALE KIDNEY CDNA, RIKEN FULL-LENGTH ENRICHED L
DE: CLONE:0610031F03, FULL INSERT SEQUENCE).
GN: S100A1.
OS: Mus musculus (Mouse).
OC: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
OC: Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Muriti
OX: NCBI_TaxID=10090;
RX: [1]
RN:
RP: SEQUENCE FROM N.A.
RC: STRAIN=C57BL/6J; TISSUE=KIDNEY;
RA: Adachi J., Aizawa K., Akhita S., Akimura T., Aono H., Ariz
```

RA Atakawa T., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanganeki T., Hara A., Hayatsu M., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kato H., Kawai J., Kojima Y.,  
RA Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T.,  
RA Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D.,  
RA Shibata K., Shibata Y., Shingawa A., Shiraki T., Sogabe Y.,  
RA Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y.,  
RA Toyota T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Morimatsu M., Hayashizaki Y.,  
RL submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.  
RN [12]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;  
RC MEDLINE=21085660; PubMed=11217851;  
RX RIKEN FANTOM Consortium.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [13]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;  
RC MEDLINE=99279253; PubMed=10349636;  
RX Carninci P., Hayashizaki Y.;  
RA "High efficiency full-length cDNA cloning.";  
RT Meth. Enzymol. 303:19-44(1999).  
RL [14]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;  
RC MEDLINE=20499374; PubMed=11042159;  
RX Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [15]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;  
RC MEDLINE=20530913; PubMed=11076861;  
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Suni N., Ishii Y., Nakamura S., Hatama M., Nishise T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [16]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAUSBERG R.;  
RL submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [17]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RA Du X.-J., Cole T.J., Tennis N., Gao X.-M., Kontgen F., Kemp B.E.,  
RA Heierhorst J.;  
RT "Impaired cardiac contractility without cardiomyopathy in S100A1-  
RT deficient mice.";  
RL submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK012578; BAB28330.1; -;  
DR EMBL: BC005590; AAH05590.1; -;  
DR EMBL: AF368423; AAL14436.1; -;  
DR EMBL: AK002721; BAB22308.1; -;  
SQ SEQUENCE 94 AA: 10505 MW: 708E817BBF36ED15 CRC64;

alignment_scores:	156.50	Length:	84
Quality:		Gaps:	
Ratio:	2.484	1	
Percent Similarity:	75.000	Percent Identity:	38.095

US-09-806-382A-2 x Q91V77 ..  
Align seg 1/1 to: Q91V77 from: 1 to: 94

[illegible]

seq\_name: sp\_human:Q9BU83

```

ID      seq_documentation_block:
ID      09BU83      PRELIMINARY:      PRT:      98 AA.
AC      Q9BU83;
DT      01-JUN-2001 (T-EMBLrel. 17, Created)
DT      01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT      01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE      S100 CALCIIUM-BINDING PROTEIN A2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=OVARY, ADENOCARCINOMA;
RA      Strausberg R.;
RL      Submitted (Feb.-2001) to the EMBL/Genbank/DBJ databases.
CC      -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR      EMBL; BC002829; AA002829.1; -.
DR      HSSP; P30801; IA03.
DR      InterPro; IPR002048; EF-hand.
DR      InterPro; IPR001751; S100_CaBP.
DR      SMART; PF01023; S.100; 1.
DR      SMART; SM00054; EFP; 1.
DR      PROSITE; PS00018; EF_HAND; UNKNOWN.1.
DR      PROSITE; PS00303; S100_CaBP; 1.
SQ      SEQUENCE 98 AA; 1117 MW; 56609548450142A9 CRC64;

```

```
alignment_scores:      Length: 87
    Quality:          155.50
    Ratio:             2.356
Percent Similarity:    75.862 Percent Identity: 36.782
```

```
alignment_block:
```

Align seg 1/1 to: Q9BU83 from: 1 to: 98

```

1  ATGACTTGC AAAATGTGCGACCTGGAAACGACACATTAAGACATCATCA  50
    ||| ||| |||:::|||||::: ::: :::::
1  MetMetCys.....SerSerLeuGluGlnAlaLeuAlaValLeuValTh  15
51  CACCTTCCACCAATATCTTGTGGAAAGCTGGGGACCCACCCACACCTGTGACC  100

```







```
alignment_scores:  
    quality: 152.50      Length: 84  
    ratio:   2.421       Gaps: 1  
Percent Similarity: 75.000 Percent Identity: 36.905
```

266 CC 267  
 85 1a 85

alignment_scores:		
Quality:	148.00	Length: 96
Ratio:	2.056	Gaps: 3
Percent Similarity:	75.000	Percent Identity: 33.333
alignment_block:		



alignment\_block:  
US-09-806-382A-2 x Q05331 ..

Align seg 1/1 to: Q05331 from: 1 to: 1218

```

13 ATGTGCAGCTGGAGACGACATAGACCATCATCAGACCTTCACCA 62
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 MetSerThrLeuValPheIleAlaIleIleAsnLeuPheAsnG1 17
63 ATACTGTGAAGCTGGGGACCCACACCCGGAACCCAGGGGGAATTC 112
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17 uTySerLysLysAspLysAsnThrAspThrLeuSerLysGluLeuL 34
113 AAGAGCTGTGCGAAAAGATCTGCAAAATTTCTCAAGAAGAGATAG 162
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 yseGlutLeuLeuGluLysGluPheArgGlnIleLeuLysAsnPro 49
163 AATGAAAAGTCTATAGACATCATGAGAGCTTGACACAAATGCAGA 212
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 AspProAspMetValAspValPheMetAspHisLeuAspHisAs 66
213 CAAGCAGCTGAGCTTGAGAGATTCATCATGCGATGCGAGGCTTA 262
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
66 nysLysIleAspPheThrGluPheLeuLeuMetValPheLysLeu 83
263 GGGCCTCCACAGAGAGATGCAC...GAGGCTGACGAGGCGCCCTG 309
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 InIaLtyrLtyrGluSerThrArgLysGluAsnLeuProIleSerG 99
310 CACCATAG 318
|||||
100 LysHisArg 102

```

seq\_name: sp\_vertibrate:Q9PSF6

seq\_documentation\_block:

```

ID Q9PSF6 PRELIMINARY; PRT; 92 AA.
AC Q9PSF6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ICAACALCIN.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
OC Siluriformes; Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9437615; PubMed-8090068;
RA Bettini E., Porta A.R., Dahmen N., Wang H., Margolis F.L.;
RT "Expressed sequence tags (EST) identify genes preferentially expressed
RT in catfish chemosensory tissues.";
RL Brain Res. Mol. Brain Res. 23:285-291(1994).
CC -1 SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC -1 SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR HSP; P30801; 1A03.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_Cabp.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Calcium-binding.
SQ SEQUENCE 92 AA; 10022 MW; E04875D0C921C50 CRC64;

```

alignment\_scores:

Quality: 140.50 Length: 90  
Ratio: 2.129 Gaps: 1  
Percent Similarity: 73.333 Percent Identity: 33.333

alignment\_block:

US-09-806-382A-2 x Q9PSF6 ..

Align seg 1/1 to: Q9PSF6 from: 1 to: 92

```

13 ATGTGCAGCTGGAGACGACATAGACCATCATCAGACCTTCACCA 62
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 MetSerAspLeuGluLysGluMetAlaLeuLeuIleSerThrPheHisL 17
63 ATACTGTGAAGCTGGGGACCCACACCCGGAACCCAGGGGGAATTC 112
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
17 stySerItyrGluGluGlyAspLysCysThrLeuThrLysGluLeuL 34
113 AAGAGCTGTGCGAAAAGATCTGCAAAATTTCTCAAGAAGAGATAG 162
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 yAspLeuLeuThrLysGluLeuGluGlyAlaPhe...GlyAsnCysSer 49
163 AATGAAAAGTCTATAGACATCATGAGAGCTTGACACAAATGCAGA 212
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 AspGlnAlaThrLeuAspLysIlePheLysAspLeuAspThrAsnAla 66
213 CAAGCAGCTGAGCTTGAGAGATTCATCATGCGATGCGAGGCTTA 262
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
66 pGlyValValAspPheGluGluItyrAlaThrMetValAlaCysThr 83
263 GGGCCTCCACAGAGAGATG 282
|||||
83 eLysCysAsnLysSerLeu 89

```

seq\_name: sp\_rudent:Q9R2B7

seq\_documentation\_block:

```

ID Q9R2B7 PRELIMINARY; PRT; 89 AA.
AC Q9R2B7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CALCYCLIN (CALCIUM BINDING PROTEIN).
GN CACY OR S100A6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-WISTAR;
RA Konrad L., Gabius H.J., Annemüller G.;
RT "Sequence and expression study of calycclin in the rat testis.";
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-WISTAR;
RA Ito M., Kizawa K.;
RT "Expression of S100 genes in hair follicle epithelium.";
RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
CC -1 SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC -1 SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AJ132717; CAB42002.1; -.
DR EMBL; AF140232; AAK28306.1; -.
DR HSP; P30801; 1A03.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_Cabp.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Calcium-binding; Cyclin.
SQ SEQUENCE 89 AA; 10035 MW; 2AA1A4163D57DC87 CRC64;

```

alignment\_scores:

Quality: 135.00 Length: 94  
Ratio: 2.213 Gaps: 2  
Percent Similarity: 64.894 Percent Identity: 31.915

alignment block:  
US-09-806-382a-2 x 09R2B7 ..

Align seg 1/1 to: 09R2B7 from: 1 to: 89

```
1 ATGACTTGCAAAATGTGGCAAGCTGGAACGCAACATAGACCATCATCAA 50
  |||::||| ::::||| ||| ::::
1 MetAlaCysProLeuAspGlnAla.....IleGlyLeuLeuValAl 14
51 CACCTTCACCAATATCTGTGAAGCTGGGGCAACCCAGACACCCCTGAAC 100
  ||||:::||||| ||| ||| ||||:::
14 allePheHisIleTyrSerGlyLysGluGlyAspLysHisThrLeuSerL 31
101 AGGGGGAATTCAAAAGAGCTGGTGGCAAAAGATCTGCAAAATTTCTCAAG 150
  :: ||:::|||||||:::||||| |||
31 yslYsgIuLeuLysGluLeuIleGlnLysGluLeuThrIleGlyAlaLys 47
151 AAGGAGATAAGATGAATGAAGGTCATAGACACATCATGTGAGAGAGCTTGA 200
  ::::: :::: ||| :::::|||||||
48 LeuGlnAspAlaGlu.....IleAlaArgLeuMetAspAspLeuAs 61
201 CACAATATGCAGACAAGCAGCTGAGCTTCGAGAGATTCATCATGTGATGG 250
  | ||| |||:::|||||:::|||||:::
61 ParGAsnLysAspGlnGluValAsnPheGlnGluTyrValAlaIaPheLeuG 78
251 CGAGGCTAACCTGGGCTCCACAGAGAGATG 282
  :: ||::: :::: ||| :::
78 LylAlaLeuAlaLeuIleTyrAsnGlnValAlaLeu 88
```

100

---

OM of: US-09-806-382a-2 to: A\_Geneseq\_032802.\* out\_format : pfs

Date: Sep 9, 2002 3:10 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=firmer.nzp.model -DEV=xlp  
-Q/cgn2.1/USPRO.spool/US09806382/runat\_09092002\_143447\_15354/app\_query.fasta\_1.752  
-DB-A\_Geneseq\_032802 -OPM=fastan -SUFFIX=rag -GAPOP=12.000  
-GAEXT=4.000 -MINMATCH=0.100 -LOOPCI=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GCAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FCAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELDP=6.000 -DELEXT=7.000 -STAR=1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pts  
-NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09806382.@CGNL\_1.120 -NCPU=6 -ICPU=3 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-806-382A-2

Query length: 345

Database: A\_Geneseq\_032802.\*

Database sequences: 747574

Database length: 111073796

Search time (sec): 103.790000

Score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU31075 +			448.00	872.50	1.1e-40
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU1760 +			417.00	815.77	2.4e-37
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU9795 +			371.00	719.20	3.1e-32
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:ABR0930 +			351.00	691.13	3.8e-30
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:ABR25051 +			351.00	691.13	3.8e-30
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU1792 +			351.00	691.13	3.8e-30
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU4591 +			351.00	691.13	3.8e-30
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU20352 +			351.00	691.13	3.8e-30
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU4612 +			351.00	691.13	3.8e-30
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU30604 +			315.50	612.82	3.6e-26
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:ABR4293 +			229.00	454.12	8.7e-17
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:ABR2578 +			229.00	454.12	8.7e-17
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU63181 +			229.00	454.12	8.7e-17
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU5955 +			229.00	454.12	8.7e-17
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU20763 +			229.00	454.12	8.7e-17
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU36103 +			229.00	454.12	8.7e-17
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:AAU3564 +			214.50	418.50	4.0e-15
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAU24137 +			214.50	418.50	4.0e-15
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAU4542 +			214.50	418.50	4.0e-15
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU31907 +			214.50	418.50	4.0e-15
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU31908 +			214.50	418.50	4.0e-15
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU30764 +			209.50	408.74	1.4e-14
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAU90765 +			206.50	402.93	3.0e-14
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:ABG7352 +			197.50	376.78	8.2e-13
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU1966 +			163.00	329.56	9.2e-10
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAU5544 +			165.50	321.57	9.7e-10
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAU5545 +			165.50	319.91	1.0e-09
/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU12007 +			164.50	320.01	1.2e-09

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU31075

seq\_documentation\_block:

ID	AAU31075 standard; Protein; 181 AA.
XX	AAU31075;
AC	XX
XX	XX
DT	18-DEC-2001 (first entry)
XX	XX
DE	Novel human secreted protein #1566.
XX	XX
KW	Human; Vaccination; gene therapy; nutritional supplement;
KW	stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW	immune suppression; immune stimulation; anti-inflammatory; Leukemia.
OS	Homo sapiens.
XX	XX
PN	WO200179449-A2.
XX	XX
PD	25-OCT-2001.
XX	XX
PF	16-APR-2001; 2001WO-US08656.
XX	XX
PR	18-APR-2000; 2000US-0552929.
PR	26-JUN-2001; 2001US-0770160.
XX	XX
PA	(HYSE-) HYSEQ INC.
PI	Tang YT, Liu C, Drmanac RT;
XX	XX
DR	WPI; 2001-611725/70.
XX	XX
PT	Nucleic acids encoding a range of human polypeptides, useful in genetic
PT	vaccination, testing and therapy -
XX	XX
PS	Claim 20; Page 400; 765pp; English.
XX	XX
CC	The invention relates to novel human secreted polypeptides. The
CC	polypeptides and antibodies to the polypeptides are useful for
CC	determining the presence of or predisposition to a disease associated
CC	with altered levels of polypeptide. The polypeptides are also useful for
CC	identifying agents (agonists and antagonists) that bind to them. Cells
CC	expressing the proteins are useful for identifying a therapeutic agent
CC	for use in treatment of a pathology related to aberrant expression or
CC	physiological interactions of the polypeptide. Vectors comprising
CC	the nucleic acids encoding the polypeptides and cells genetically
CC	engineered to express them are also useful for producing the proteins.
CC	The proteins are useful in genetic vaccination, testing and
CC	therapy, and can be used as nutritional supplements. They may be used to
CC	increase stem cell proliferation; to regulate haematopoiesis; and in
CC	bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC	immune suppression and/or stimulation; as anti-inflammatory agents; and
CC	in treatment of leukemias. AAU29510-AAU3304 represent the amino acid
CC	sequences of novel human secreted proteins of the invention.
XX	XX
SQ	Sequence 181 AA:
alignment_scores:	
Quality:	448.00 Length: 114
Ratio:	4.480 Gaps: 3
Percent Similarity:	87.719 Percent Identity: 82.456
alignment_block:	
US-09-806-382A-2 x AAU31075	..
Align seg 1/1 to: AAU31075 from: 1 to: 181	
13 ATGCGCAGCGTGAAGCAACATAGACATCAACACCTTCACCA 62	
4 MetcrlpImleuclInlrpAsnllleGluThrIleIeAsnThrPheHis61 20	
63 ATACTGTGTAAGCTGGGCGACCCAGACACCCCTGAACAGGGGGAATTCA 112	
20 nTyrsetVallyslleuclYHisrproAsphrIleuAsnGlnGlyIuphel 37	

```

113 AAGACCTGGTGGGAAAAGATCTG...CAAAATTTCTCAAGAGAGAAAT 159
      |||
37 YSGILNValAlarGlyAspLeuGlyGlnAspPheLeuYslySGIuaSn 53
      |||
160 AAGATGAAAAGGTATAGACACATCATGAGAGACCTGGACACAATCC 209
      |||
54 LysAsnGluLysValIleGlnHisIleHisGluAspLeuAspThrAsnAl 70
      |||
210 AAGACAGCAGCTGAGCTTCGAGAGATTCATGCTGATGGAGAGGCTTAA 259
      |||
70 AAlaIleGlnLeuSerPheGluGluPheIleMetLeuMetAlaAlaGAlaL 87
      |||
260 CCTGGCCCTCC...CAGCAGAAAGATGCACAGAGGTGACAGAGGCCCT 303
      |||
87 YSPROGlyAlaLeuProThrArgArgMetHisGluGlyAspIysGlyPro 103
      |||
304 ...GGCCACACCATTAAGCCAGGCTCGGGGAGGACCCCC 342
      |||
104 ThrProProProHisLysProGlyLeuGlyGlyThrPro 117

seq_name: /SIDSL/gcdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AA41780

seq_documentation_block:
ID   AAM41780 standard; Protein; 114 AA.
XX
AC   AAM41780;
XX
DE   22-OCT-2001 (first entry)
XX
DE   Human polypeptide SEQ ID NO 6711.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PE   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI   Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR   WPI: 2001-442253/47.
DR   N-PSDB: AAI60936.
XX
DE   Novel nucleic acids and polypeptides, useful for treating disorders
XX   such as central nervous system injuries -
XX   Example 2: SEQ ID NO 6711; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AA157798-AA161369) and
CC   the encoded polypeptides (AAM38642-AA4213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide

```

```

CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/Inhibin activity, chemotactic/chemokine activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemia and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 114 AA;
XX

alignment_scores:
      Quality: 417.00      Length: 108
      Ratio: 4.299      Gaps: 0
      Percent Similarity: 89.815      Percent Identity: 71.296

alignment_block:
US-09-806-382A-2 x AAM41780
Align seg 1/1 to: AAM41780 from: 1 to: 114

1 ATGACTTGCAGAAATGTCGAGCTGAGACCAACATAGACCATCATCA 50
  |||
7 MetThrCylLysMetSerGlnLeuGlnArgAsnIle***ThreIleAs 23
  |||
51 CACCTTCACCAATACTCTGTGAAGCTGGGACCCAGACACCTTAACC 100
  |||
23 nTrIleuHisHisIstYSerValLysLeuGlnHisProAspThrLeuIleH 40
  |||
101 AGGGGAATTCAGAAAGCTGTGTCGAAAGATCTGCAAAATTTCTCAAG 150
  |||
40 isoIyluPheLysGluLeuValArgThrAspLeuHisnIleLeuMet 56
  |||
151 AAGGAGATTAAGATGAAGAGTCAATAGAACACATCATGAGGACCTGGA 200
  |||
57 LysGluAsnLysAsnAspGlnAlaIle***HisIleMetGluAspLeuAs 73
  |||
201 CACAATTCAGACACAGCTGAGCTTCGAGAGTTCATCATGCTGATG 250
  |||
73 pThrAsnAlaHisMetGlnIleIlePheLysGluLeuIleMetLeuMet 90
  |||
251 CGAGCTTAACCTGGGCTCCACAGACAGATGCAGAGGTGACGAGGCG 300
  |||
90 lAmetLeuThrTrpSerTyRHIsAspAsnMetHisAspAlaAspTyrgly 106
  |||
301 CCTGGCCACCAACCATTAAGCCAGGC 324
  |||
107 ProGlyGlnGlnHisArgProGly 114

seq_name: /SIDSL/gcdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAU29795

seq_documentation_block:
ID   AAU29795 standard; Protein; 211 AA.
XX
AC   AAU29795;
XX
DE   18-DEC-2001 (first entry)
XX
DE   Novel human secreted protein #286.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX

```





```
1 LysGluAsnLysAsnGlnLysValIleGlnHisIleMetGluAspLeuAs 17
201 CACAATGCGAGACAAGCAGCTGAGCTTCGAGAGATTATCATGTCGATGG 250
17 pThrAsnAlaAspLysGlnLeuSerPheGlnLupPheIleMetLeuMet 34
251 CGAGGCTTAACCTGGGCTCCCGACGAGATGCACGAGGGTGACGAGGCG 300
34 laArgLeuThrTrpAlaSerHisGlnLysMetHisGlnLupSpGlnGly 50
301 CCTGGCCACACCATTAAGCAGGCTCTGGGAGGAGGCCACCCCG 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGlnGlyThrPro 64
seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:ABB25051
seq_documentation_block:
ID ABB25051 standard; Protein: 64 AA.
AC ABB25051;
XX
XX 23-JAN-2002 (first entry)
DT
DE Protein #7050 encoded by probe for measuring heart cell gene expression.
XX
XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX
XX Homo sapiens.
OS
XX WO200157274-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00666.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488899/53.
DR
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
PS
XX Claim 15; SEQ ID No 26821; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 64 AA:
64
alignment_scores:
6
4
```

```
Quality: 351.00 Length: 64
Ratio: 5.484 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-806-382A-2 x ABB25051 ..
Align seg 1/1 to: ABB25051 from: 1 to: 64
151 AAGAGATTAAGATGAAAAGTGCATAGAACACATATGAGGAGCTGGA 200
|||||
1 LysGluAsnLysAsnGlnLysValIleGlnHisIleMetGluAspLeuAs 17
17 pThrAsnAlaAspLysGlnLeuSerPheGlnLupPheIleMetLeuMet 34
201 CACAATGCGAGACAAGCAGCTGAGCTTCGAGAGATTATCATGTCGATGG 250
|||||
17 pThrAsnAlaAspLysGlnLeuSerPheGlnLupPheIleMetLeuMet 34
251 CGAGGCTTAACCTGGGCTCCCGACGAGATGCACGAGGGTGACGAGGCG 300
|||||
34 laArgLeuThrTrpAlaSerHisGlnLysMetHisGlnLupSpGlnGly 50
301 CCTGGCCACACCATTAAGCAGGCTCTGGGAGGAGGCCACCCCG 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGlnGlyThrPro 64
seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:AAM61792
seq_documentation_block:
ID AAM61792 standard; Protein: 64 AA.
XX
XX AAM61792;
AC
XX
XX 05-NOV-2001 (first entry)
DT
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 33897.
DE
XX
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KM epilepsy; cancer.
XX
XX Homo sapiens.
OS
XX WO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00667.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-483446/52.
DR
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
PS
XX Example 4; SEQ ID NO: 33897; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
```

CC the probes of the invention.

XX Sequence 64 AA;

alignment\_scores:

Quality:	351.00	Length:	64
Ratio:	5.484	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-806-382A-2 x AAM61792 ..

Align seg 1/1 to: AAM61792 from: 1 to: 64

```
151 AAGGAGATTAAGATGAAGAAGTCATAGAACATCATGAGAGACCTGGA 200
|||||
1  LysGluAsnLysAsnGluLysValIleGluHisIleMetGluAspLeuAs 17
201 CACAAATGCAGACAGCAGCTGAGCTTCGAGAGATTGATCATGCTGATGG 250
|||||
17 pPrrAsnAlaAspLysGlnLeuSerPheGluGluPheIleMetLeuMetA 34
251 CGAGGCTTAACCTGGGCTCCCGACGAGAATGCACGAGGCTGACGAGGCC 300
|||||
34 IaArgLeuThrTyrPalaSerHisGluLysMetHisGluGlyAspGluGly 50
301 CCGGGCCACACCATTAAGCCAGGCTCGGGGAGGAGGCCCC 342
|||||
51 ProGlyHisHisHisHisLysProGlyLeuGlyGluGlyThrPro 64
```

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT: AAM74591

seq\_documentation\_block:

ID AAM74591 standard; Protein; 64 AA.

AC AAM74591;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34897.

KW Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; Leukaemia; Lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00668.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 34897; 658bp + Sequence Listing; English.

PS The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention.

XX Sequence 64 AA;

Quality:	351.00	Length:	64
Ratio:	5.484	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-806-382A-2 x AAM74591 ..

Align seg 1/1 to: AAM74591 from: 1 to: 64

```
151 AAGGAGATTAAGATGAAGAAGTCATAGAACATCATGAGAGACCTGGA 200
|||||
1  LysGluAsnLysAsnGluLysValIleGluHisIleMetGluAspLeuAs 17
201 CACAAATGCAGACAGCAGCTGAGCTTCGAGAGATTGATCATGCTGATGG 250
|||||
17 pPrrAsnAlaAspLysGlnLeuSerPheGluGluPheIleMetLeuMetA 34
251 CGAGGCTTAACCTGGGCTCCCGACGAGAATGCACGAGGCTGACGAGGCC 300
|||||
34 IaArgLeuThrTyrPalaSerHisGluLysMetHisGluGlyAspGluGly 50
301 CCGGGCCACACCATTAAGCCAGGCTCGGGGAGGAGGCCCC 342
|||||
51 ProGlyHisHisHisHisLysProGlyLeuGlyGluGlyThrPro 64
```

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT: AAM20352

seq\_documentation\_block:

ID AAM20352 standard; Protein; 64 AA.

AC AAM20352;

DT 12-OCT-2001 (first entry)

DE Peptide #6786 encoded by probe for measuring cervical gene expression.

KW Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer.

OS Homo sapiens.

PN WO200157278-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00670.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

XX Example 4; SEQ ID NO: 34897; 658bp + Sequence Listing; English.

PS The present invention provides a number of single exon nucleic acid

```
XX Claim 27; SEQ ID No 25178; 487bp; English.
PS
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AA110068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 64 AA;

alignment_scores:
Quality: 351.00 Length: 64
Ratio: 5.484 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-806-382A-2 x AAM20352 ..
Align seg 1/1 to: AAM20352 from: 1 to: 64

151 AAGGAGATATAGATAAGTGAAGGTCATAGACATCATGAGGAGGAGCTGGA 200
|||||
1 LysGluAsnLysAsnGluLysValIleGluHisIleMetGluAspLeuAs 17
201 CACAAATGCAGACAGACAGCTGAGCTTCAGAGAGTTTCATCATGCTGATGG 250
|||||
17 pThrAsnAlaAspLysGlnLeuSerPheGluGluPheIleMetLeuMetA 34
251 CGAGGCTAACCTGGGCTCCCGACAGAGATGCACGAGGCTGACGAGGC 300
|||||
34 IaArgLeuThrTrpAlaSerHisGlnLysMetHisGlnGlyAspGluGly 50
301 CCTGGCCACACCATTAAGCCAGGCTCGGGAGGAGGACACCCG 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGluGlyThrPro 64

seq_name: /SIDSI/gcgdata/hold_geneseq/geneseq-emb1/AA2001.DAT:AAM34703
seq_documentation_block:
ID AAM34703 standard; Protein: 64 AA.
XX
AC AAM34703;
XX
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #8740 encoded by probe for measuring placental gene expression.
XX
KW Probe: microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001MO-US00663.
XX
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
```

```
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID No 34972; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AA131315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 64 AA;

alignment_scores:
Quality: 351.00 Length: 64
Ratio: 5.484 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-806-382A-2 x AAM34703 ..
Align seg 1/1 to: AAM34703 from: 1 to: 64

151 AAGGAGATATAGATAAGTGAAGGTCATAGACATCATGAGGAGGAGCTGGA 200
|||||
1 LysGluAsnLysAsnGluLysValIleGluHisIleMetGluAspLeuAs 17
201 CACAAATGCAGACAGACAGCTGAGCTTCAGAGAGTTTCATCATGCTGATGG 250
|||||
17 pThrAsnAlaAspLysGlnLeuSerPheGluGluPheIleMetLeuMetA 34
251 CGAGGCTAACCTGGGCTCCCGACAGAGATGCACGAGGCTGACGAGGC 300
|||||
34 IaArgLeuThrTrpAlaSerHisGlnLysMetHisGlnGlyAspGluGly 50
301 CCTGGCCACACCATTAAGCCAGGCTCGGGAGGAGGACACCCG 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGluGlyThrPro 64

seq_name: /SIDSI/gcgdata/hold_geneseq/geneseq-emb1/AA2001.DAT:ABB44612
seq_documentation_block:
ID ABB44612 standard; Protein: 113 AA.
XX
XX
AC ABB44612;
XX
XX
DT 25-JAN-2002 (first entry)
XX
DE Mouse wound healing related polypeptide SEQ ID NO 101.
XX
KW Human; mouse; vulnery; dermatological; skin disorder; wound healing;
KW gene therapy.
XX
OS Mus musculus.
XX
XX
PN CA2325226-A1.
XX
XX
PD 17-MAY-2001.
XX
XX
PR 16-NOV-2000; 2000CA-2325226.
XX
XX
PR 17-NOV-1999; 99DE-1055349.
PR 17-DEC-1999; 99US-0172511.
PR 20-JUN-2000; 2000DE-1030149.
```

XX (SWIT-) SWITCH BIOTECH AG.  
 XX Regenbogen J, Wolf E, Goppelt A, Werner S, Halle J;  
 XX MPI, 2001-433142/47.  
 XX  
 XX Use of novel polypeptide or its variant or nucleic acid encoding the  
 XX polypeptide for diagnosing and/or preventing and/or treating skin  
 XX disorders and/or treatment in wound healing or for identifying active  
 XX substances -  
 XX  
 XX Claim 3; Page 245; 265pp; English.  
 XX  
 XX The invention relates to the use of a polypeptide (ABBA4544-ABBA4601,  
 XX ABBA4606-ABBA4623) or its variant or encoding nucleic acid  
 XX (ABBA81990-ABBA81995, ABBA82016-ABBA82032) with vulnerary and/or  
 XX dermatological activity for the diagnosis, prevention and treatment of  
 XX skin disorders and treatment in wound healing or for the identification  
 XX of pharmacologically active substances. The nucleic acids are useful in  
 XX gene therapy.  
 XX Note: The printed sequence listing for this specification was incomplete,  
 XX terminating part way through SEQ ID NO 106. The remaining data was  
 XX obtained from EPO data for an equivalent patent (EP1114862).  
 XX  
 XX Sequence 113 AA;

alignment\_scores:      Quality: 329.50      Length: 107  
 Ratio: 3.702      Gaps: 1  
 Percent Similarity: 83.178      Percent Identity: 59.813

alignment\_block:

US-09-806-382A-2 x ABBA4612 ..

Align seg 1/1 to: ABBA4612 from: 1 to: 113

```

16 TCGCAGCTGGAAGCAGACATCATGACCATCAACACCTTCACCAATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7 SerGImetGluArgSerIleThrIleIleAspThrPheHisGlnIty 23
66 CUCGTGGAAGCTGGGGCAGCCAGACACCTGAGACAGGGGGAATCAAG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
23 rterArGlySGluGlyHisProAspThrLeuSerLysGluPheArG 40
116 AGCTGTCGGAAGAGTGTGCAAAATTTCTCAAGAGAGAGATAGAAT 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
40 ImetValGluAlaGlnLeuAlaThrPheMetLysGluLysArgGsn 56
166 GAAAGGTCATAGAACACATCATGAGACCTGAGACACAATGACAGCAA 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
57 GluAlaLeuIleAsnAspIleMetGluAspThrAsnGlnAspAs 73
216 GGAGCTGAGCTTCGAGAGTTCATCATGCTGAGGAGGCTAACCTGGG 265
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
73 ngInLeuSerPheGluGluCysMetMetLeuAlaLysLeuIlePheA 90
266 CCTCCACGAGAGATGACAGAGGCTGAC...GAGGCGCTGAGCCACAC 312
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
90 lacYHisGluLysLeuHisGluAsnAsnProArgGlyHisLysHisSer 106
313 CATAGCCAGGCGCTGGGGAG 333
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
107 HisGlyLysGlyCysGlyLys 113

```

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:AAU30604

seq\_documentation\_block:

ID AAU30604 standard; Protein: 153 AA.

XX AC AAU30604;

XX

DT 18-DEC-2001 (first entry)  
 XX  
 XX Novel human secreted protein #1095.  
 DE  
 XX  
 XX Human: vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX MO200179449-A2.  
 PN  
 XX  
 XX 25-OCT-2001.  
 PD  
 XX  
 XX 16-APR-2001; 2001WO-US08656.  
 PF  
 XX  
 XX 18-APR-2000; 2000US-0552929.  
 PR  
 XX 26-JAN-2001; 2001US-0770160.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 PI  
 XX  
 XX MPI, 2001-611725/70.  
 DR  
 XX  
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy -  
 PT  
 XX  
 XX  
 XX Claim 20; Page 317; 765pp; English.  
 PS  
 XX  
 XX The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising  
 CC the nucleic acids encoding the polypeptides and cells genetically  
 CC engineered to express them are also useful for producing the proteins.  
 CC The proteins are useful in genetic vaccination, testing and  
 CC therapy, and can be used as nutritional supplements. They may be used to  
 CC increase stem cell proliferation; to regulate haematopoiesis; and in  
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
 CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid  
 CC sequences of novel human secreted proteins of the invention.  
 XX  
 XX Sequence 153 AA;

alignment\_scores:      Quality: 315.50      Length: 122  
 Ratio: 3.286      Gaps: 7  
 Percent Similarity: 78.689      Percent Identity: 63.115

alignment\_block:

US-09-806-382A-2 x AAU30604 ..

Align seg 1/1 to: AAU30604 from: 1 to: 153

```

1 ATGACTTGCAAAATGTGGCAG...CTGGAACGCAACATAGACCATCAT 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
32 MetThrCysGlyMetProGlnHisValThrGlnIn**ArgProIleIl 48
48 CAACACCTTCACCAATATCTGTGGAAGCTGGGGCAGCCAGACACCTGA 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
48 eAsnThrSerHisGlnIleValLysLeuGlyHisProAspThrIleuA 65
98 ACCAGGGGAATTCAAAGAGCTGTGCGAAGAGATCTGCAAAATTTTCTC 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
65 sngInGlyGluPheLysGluLeuValArgLysAspLeuGlnAsnPheLeu 81

```

```
148 AAGAGGAGATATAGATGAAAAGTCAATGACATCATGAGGAGCCT 197
|||||
82 LysLysGLAsnLysAsnGluLysValIIEgluHISILeMetArgLysPr 98
198 GGACACA...AATGCAGACAAG.....CAGCTGAGCTTCGAGGAGTCA 238
||| ||||| ::| ::| |||||
98 oGlyThrGlnAsnAlaGlnProAlaGluLeuSer...ArgGluPheI 114
239 TCATGCTGATGGGAGGCTAACCTGGGCGCTCC.....CACGAGAGATG 282
|||||
114 LeMetLeuMetGlyGluAla**ProGlyAlaPheProArgLysTle 130
283 CACGAGAGGT...GACGAGGCGCCTGGCCACACCATATAG...CCAGGCGT 326
::|::| ||||| ||||| |||||
131 AlaArgGlyLeuThrGluGlyProGlyHisProAsnLys**ProGlyPr 147
327 CGGGAGGGGACACCC 342
||| |||||
147 oGlyGlyGlyAlaPro 152

seq_name: /SIBS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:AAU29794
seq_documentation_block:
ID AAU29794 standard; Protein; 57 AA.
XX
AC AAU29794;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #285.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HXSE-) HXSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 195; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
```

```
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 57 AA;

alignment_scores:
Quality: 267.00 Length: 56
Ratio: 4.944 Gaps: 0
Percent Similarity: 96.429 Percent Identity: 85.714

alignment_block:
US-09-806-382A-2/rev x AAU29794 ..
Align seg 1/1 to: AAU29794 from: 1 to: 57

188 ATGATGTGTTCTATGACCTTTTCATCTTATTCCTCTTGAGAAAT 139
|||||
1 MetMetCysSerMetThrLeuSerPheIlePheSerPheMetArgLysIle 17
138 TTGCAGATCTTTGCGACCCAGCTCTTGAATTCCTGGTGGAGGTGT 89
:|||||
17 ucysArgSerIleArgAlaSerSerTrpAsnSerProTrpPheArgValS 34
88 CTGGGTGCCCCACAGCTTCACAGAGTATTGTGGAAGGTGTGATGAGGTC 39
|||||
34 exGlyCysProSerPheThrGluTyrTrpTrpLysValLeuMetMetVal 50
38 TCTATGTTGCGTTCCAGC 21
|||||
51 TyrMetLeuArgSerSer 56

seq_name: /SIBS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:ABB42295
seq_documentation_block:
ID ABB42295 standard; Peptide; 44 AA.
XX
AC ABB42295;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #9801 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234667.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human fetal liver -
XX
PS Claim 27; SEQ ID NO 34930; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
```

CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 44 AA;

alignment\_scores:

Quality: 229.00 Length: 44  
Ratio: 5.205 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-806-382A-2 x ABB42295 ..

Align seg 1/1 to: ABB42295 from: 1 to: 44

19 CAGCTGAACGCAACATAGACCATCATCAACACCTTCACCAATATCTC 68  
|||||  
1 GlnleuGlutArgAsnIleGluThrIleLeuSnrhrPheHisGlnTyrSe 17  
69 TGTGAAGCTGGGGCAACCCAGACCCCTGAACCGAGGGGAATTCAAGAGC 118  
|||||  
17 TVallyslLeuGlyHisProAspThrLeuAsnGlnGlyIuphLeysGluL 34  
119 TGTGCGCAAGAAGATCTGCAGAAATTTCTCAG 150  
|||||  
34 euValArgLyAspLeuGlnAsnPhLeuLyS 44

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:ABB25798

seq\_documentation\_block:

ID ABB25798 standard; Protein: 44 AA.

XX ABB25798;

XX 23-JAN-2002 (first entry)

DE Protein #7797 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR.

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX

PS Claim 15; SEQ ID No 27568; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABM41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 44 AA;

alignment\_scores:

Quality: 229.00 Length: 44  
Ratio: 5.205 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-806-382A-2 x ABB25798 ..

Align seg 1/1 to: ABB25798 from: 1 to: 44

19 CAGCTGAACGCAACATAGACCATCATCAACACCTTCACCAATATCTC 68  
|||||  
1 GlnleuGlutArgAsnIleGluThrIleLeuSnrhrPheHisGlnTyrSe 17  
69 TGTGAAGCTGGGGCAACCCAGACCCCTGAACCGAGGGGAATTCAAGAGC 118  
|||||  
17 TVallyslLeuGlyHisProAspThrLeuAsnGlnGlyIuphLeysGluL 34  
119 TGTGCGCAAGAAGATCTGCAGAAATTTCTCAG 150  
|||||  
34 euValArgLyAspLeuGlnAsnPhLeuLyS 44

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AA63181

seq\_documentation\_block:

ID AA63181 standard; Protein: 44 AA.

XX AA63181;

XX 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35286.

XX Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX

PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -

XX  
PS Example 4; SEQ ID NO: 35286; 650pp + Sequence Listing; English.  
XX

CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.

XX  
SQ Sequence 44 AA;

alignment\_scores:  
Quality: 229.00 Length: 44  
Ratio: 5.205 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-806-382A-2 x AAM63181 ..

Align seg 1/1 to: AAM63181 from: 1 to: 44

19 CAGCTGGAACGCACATAGAGACATCATCAACACCTTCACCAATACTC 68  
|||||  
1 GlnLeuGluArgAsnIleGluThrIleAsnThrPheHisGlnTyrse 17  
69 TGTGAAGCTGGGGCACCCAGACACCTTGACACGGGGGATTCAGAGAC 118  
|||||  
17 rValIysLeuGlyHisProAspThrLeuAsnGlnIleGlyIuPheLysGluL 34  
119 TGTGCGAAGAGATCTGCAGAAATTTCTCAAG 150  
|||||  
34 euValArgLysAspLeuGlnAsnPheLeuLys 44



OM of: US-09-806-382a-2 to: Pending\_Patents\_AA\_Main:\* out\_format : pfs  
Date: Sep 9, 2002 3:22 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

```
-MODEL=frim+np.model -DEV=xlp  
-O/cgn2.1/USPRO.spool/US09806382/rnat.09092002.143848.15457/app-query.fasta.1.752  
DB=Pending_Patents_AA_Main -QPM=fastan -SUFFIX=rapm  
-GAPOP=12.000 -GAPOP=4.000 -MINMAP=0.100 -LOOPEL=0.000  
-LOOPEXT=0.000 -GAPOP=4.500 -GAPOP=10.000  
-YGAPEXT=0.500 -YGAPOP=6.000 -YGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=bloms62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200  
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pts -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US09806382.@CGN1.1.571 -NCPU=6 -ICPU=3 -LONLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY -WAIT -THREADS=1
```

## Search information block:

```
Query: US-09-806-382a-2  
Query length: 345  
Database: Pending_Patents_AA_Main:*  
Database sequences: 3502263  
Database length: 351980561  
Search time (sec): 409.890000
```

## score list:

```
Sequence Strd Orig ZScore EScore Len 1 Documentation  
/cgn2.6/ptodata/2/paa/PCtus.COMB.pcp:US-09-80656-7634 + 448.00 808.16 1.3e-36 18  
/cgn2.6/ptodata/2/paa/US094.COMB.pcp:US-09-488-725A-6711 + 417.00 755.97 1.6e-33 11  
/cgn2.6/ptodata/2/paa/PCtus.COMB.pcp:US-09-488-725A-6711 + 371.00 666.54 8.5e-29 21  
/cgn2.6/ptodata/2/paa/US087.COMB.pcp:US-08-759-913-7 + 335.50 644.06 2.8e-27 113  
/cgn2.6/ptodata/2/paa/US092.COMB.pcp:US-09-205-680-7 + 335.50 644.06 2.8e-27 113  
/cgn2.6/ptodata/2/paa/US094.COMB.pcp:US-09-492-026-7 + 351.00 641.09 7.3e-27 64  
/cgn2.6/ptodata/2/paa/PCtus.COMB.pcp:US-09-864-761-40349 + 351.00 641.09 7.3e-27 64  
/cgn2.6/ptodata/2/paa/US098.COMB.pcp:US-09-714-593-101 + 329.50 596.71 1.2e-24 113  
/cgn2.6/ptodata/2/paa/US097.COMB.pcp:US-09-714-593-101 + 329.50 596.71 1.2e-24 113  
/cgn2.6/ptodata/2/paa/US060.COMB.pcp:US-60-340-187-1029 + 320.00 579.58 1.1e-23 111  
/cgn2.6/ptodata/2/paa/US060.COMB.pcp:US-60-340-187-1030 + 320.00 579.58 1.1e-23 111  
/cgn2.6/ptodata/2/paa/US060.COMB.pcp:US-60-340-187-1031 + 320.00 579.58 1.1e-23 111  
/cgn2.6/ptodata/2/paa/PCtus.COMB.pcp:US-09-834-366-20109 + 289.00 530.10 1.4e-20 52  
/cgn2.6/ptodata/2/paa/US098.COMB.pcp:US-09-834-366-20109 + 289.00 530.10 1.4e-20 52  
/cgn2.6/ptodata/2/paa/PCtus.COMB.pcp:US-09-834-366-20109 + 289.00 530.10 1.4e-20 52  
/cgn2.6/ptodata/2/paa/US060.COMB.pcp:US-60-197-873-20109 + 267.00 489.36 2.4e-18 56  
/cgn2.6/ptodata/2/paa/PCtus.COMB.pcp:US-09-80656-7634 + 229.00 422.38 1.6e-14 4
```

```
seq_name: /cgn2.6/ptodata/2/paa/PCtus.COMB.pcp:US-09-80656-7634  
seq_documentation_block:  
; Sequence 7634, Application PC/TUS0108656  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 2172-066  
; CURRENT APPLICATION NUMBER: PCT/US01/08656  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 09/522,929  
; PRIOR FILING DATE: 2000-04-18  
; PRIOR APPLICATION NUMBER: 09/770,160  
; PRIOR FILING DATE: 2001-01-26  
; NUMBER OF SEQ ID NOS: 10994  
; SOFTWARE: Custom  
; SEQ ID NO 7634  
; LENGTH: 181  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN
```

```
; LOCATION: (6)..(43)  
; OTHER INFORMATION: S-100/ICABP type calcium binding protein domain identified by  
; OTHER INFORMATION: EMATRIX, accession number BL00303A, p-value=6.667e-26, raw sco  
; OTHER INFORMATION: of 21.77  
; NAME/KEY: DOMAIN  
; LOCATION: (7)..(55)  
; OTHER INFORMATION: S-100/ICABP type calcium binding domain identified by Pfam,  
; OTHER INFORMATION: accession name S_100, E-value=2e-08, Pfam score of 41.4  
PCT-US01-08656-7634
```

## alignment\_scores:

Quality:	448.00	Length:	114
Ratio:	4.480	Gaps:	3
Percent Similarity:	87.719	Percent Identity:	82.456

## alignment\_block:

US-09-806-382a-2 x PCT-US01-08656-7634 ..

Align seg 1/1 to: PCT-US01-08656-7634 from: 1 to: 181

```
13 ATGTGCGACGTGGAAGCAACATGACATCATCAACACTTCCACCA 62  
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
4 MettrpGlnLeuGluArgAsnIleGluThrIleIleSnrThrPheHsGl 20  
63 ATACTGTGTAAGCTGGGGCACCAGACACCTTGACACGGGGAATCA 112  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
20 ntYserValIysLeuGlyHisProAspThrLeuAsnGlnGlyIubHel 37  
113 AAGAGCTGTGCGAAGAAATCTG...CAAAATTTTCTCAGAAGCAAT 159  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
37 YsgIuLeuValArgLysAspLeuGlyGlnAsnPheLeuLysLysGluAsn 53  
160 AAGATGAAGAGTCATATGACATCATCATGAGGAGCTGACACAATGC 209  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
54 LysAsnGlnLysValIleGlnHisIleHisLysLysAspThrAsnAl 70  
210 AGACAAGAGCTGACCTTGAAGAGATTCATGCTGATGGCGAGCTAA 259  
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
70 AlaGlnGlnLeuSerPheGlnGluPheIleMetLeuMetLysAlaAl 87  
260 CCTGGGCGCTGC.....CAGCAAGATGCAGGAGGAGGCGGCGCT 303  
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
87 YspProGlyAlaLeuProThrArgAlaGlyMetHisGlyLysPlyGlyPro 103  
304 ...GGCCACCAACATATACCAAGCGCTCGGGAGGGGACACCC 342  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
104 TrpProProHisLysPheGlyLysGlyGlyThrPro 117  
seq_name: /cgn2.6/ptodata/2/paa/US094.COMB.pcp:US-09-488-725A-6711  
seq_documentation_block:  
; Sequence 6711, Application US/09488725A  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides  
; FILE REFERENCE: 784FLPCT  
; CURRENT APPLICATION NUMBER: US/09/488,725A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US/09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: US09/598,042  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: US09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: US09/653,450  
; PRIOR FILING DATE: 2000-08-31  
; PRIOR APPLICATION NUMBER: US09/662,191  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: US09/693,036  
; PRIOR FILING DATE: 2000-10-19
```

```
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 6711
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(114)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-09-488-725A-6711
```

```
alignment_scores:
  Quality: 417.00      Length: 108
  Ratio: 4.299        Caps: 0
  Percent Similarity: 89.815      Percent Identity: 71.296
```

```
alignment_block:
US-09-806-382A-2 x US-09-488-725A-6711 ..
```

Align seg 1/1 to: US-09-488-725A-6711 from: 1 to: 114

```
1 ATGACTTGCAAAATGTCGACGTGAGACGCAACATAGAGACCATCATCAA 50
|||||
7 MetThrCysLysMetSerGlnLeuGlnArgAsnIle**ThiMetIleAs 23
51 CACCTTCACCAATACTCTGTGAGCTGGGGACCCAGACACCTGGAACC 100
|||||
23 nThreuthSHstYrSerValLysLeuGlnIstProAspThrIleIleH 40
101 AGGGGGAATTCAAAGAGCTGTGTGCAAAAAGATCTGCAAAATTTCTCAAG 150
|||||
40 IsglyGlnPheLysGlnLeuValArgThrAspLeuHisAsnIleLeuMet 56
151 AAGGGGAATAGAAATGATGAAGGTCTAGACACATCATGAGGAGACTGGA 200
|||||
57 LysGlnAsnLysAsnAspGlnIleIle**HisIleMetGlnAspLeuAs 73
201 CACAATATCAGACAGACAGCTGAGCTTCGAGAGATTTCATCATGCTGATGG 250
|||||
73 PThrAsnIleHisMetGlnIleIlePheLysGlnLeuIleMetLeuMetA 90
251 CGAGGCTTAACCTGGCCCTCCACAGAGAAGATGACAGAGGTGACGAGGCG 300
|||||
90 IamLeuThrTyrSerTyrHisAspAsnMetHisAspAlaAspTyrGly 106
301 CCTGGCCACGACCATACGACGAGC 324
|||||
107 ProGlyGlnGlnHisArgProGly 114
```

seq\_name: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep:PCT-US01-08656-5822

```
seq_documentation_block:
; Sequence 5822, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 5822
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: DOMAIN
; LOCATION: (87)..(124)
; OTHER INFORMATION: S-100/ICaBP type calcium binding protein domain identified by
; OTHER INFORMATION: EMATRIX, accession number B100303B, p-value=2.698e-13, raw sco
; OTHER INFORMATION: of 26.15
; NAME/KEY: DOMAIN
; LOCATION: (36)..(85)
; OTHER INFORMATION: S-100/ICaBP type calcium binding domain identified by Pfam,
; OTHER INFORMATION: accession name S_100, E-value=0.094, Pfam score of 7.3
; NAME/KEY: misc_feature
; LOCATION: (1)...(211)
; OTHER INFORMATION: Xaa = x or * as defined in Table 2
PCT-US01-08656-5822
```

```
alignment_scores:
  Quality: 371.00      Length: 120
  Ratio: 3.747        Caps: 5
  Percent Similarity: 82.500      Percent Identity: 68.333
```

```
alignment_block:
US-09-806-382A-2 x PCT-US01-08656-5822 ..
```

Align seg 1/1 to: PCT-US01-08656-5822 from: 1 to: 211

```
1 ATGACTTGCAAAATGTCGACG...CTGGAACGCAACATAGAGACCATCAT 47
|||||
32 MetThrCysGlyMetSerProGlnHisValThrGlnGln**ArgProIleI 48
48 CAACACCTTCACCAATACTCTGTGAAAGCTGGGGACCCA...GACACCC 94
|||||
48 easnThrSerHisGlnTyrSerValLysLeuGlnIstProArgHisProG 65
95 TGACACGAGGGGAATTCAAAGAGCTGTGCGAAAGATGTCGAAATTTT 144
|||||
65 LuthrArgGlyArgPheLysGlnLeuValArgLysAspLeuIlnsnPhe 81
81 LuthrArgGlyArgPheLysGlnLeuValArgLysAspLeuIlnsnPhe 81
82 LeuLysLysGlnAsnLysAsnGlnLysValIleGlnHisIleMetGlnAs 98
145 CTCAAGAGAGAGAAATAGAAATGAAAGGTCTATAGAACATCATGAGAGA 194
|||||
82 LeuLysLysGlnAsnLysAsnGlnLysValIleGlnHisIleMetGlnAs 98
195 CCTGACACCAATATGACAGACACAGCTGAGCTTC...GAGGAGTTTCATCA 241
|||||
98 PLeuAspThrAsnAlaAspLysGlnLeuSerPheArgGlnGlnPheIleI 115
242 TGCTGATGGCGAGGCTTAACCTGGGCTCC...CAGGAGAAGATGGC 285
|||||
115 eLeuMetGlyGlnAla**ProGlyAlaPheProArgArgLysIleAla 131
286 GAGGCT...GACGAGGGCCCTGGCCACACCATTAAGCCAGGCTCGGGGA 332
|||||
132 ArgGlyLeuThrGlnGlyProGlyHisHisLysLysProGlyProGlyG 148
333 GGGCACCCCC 342
|||||
148 yGlyAlaPro 151
```

seq\_name: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:US-08-759-913-7

```
seq_documentation_block:
; Sequence 7, Application US/08759913
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN S100 PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
```



```

216 GCAGCTGACCTTCGAGAGTTCATGCTGATGCGAGGCTAACCTGGG 265
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73 nglnleuserpneglucLucysmetleuMetglyLysleuIlephe 90
266 CCTCCACAGAGAAGATGACAGAGGTGACGAGGCCCTGGCCACACCAT 315
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90 lacyshtsluLysleuHtSgluasnAhnPro...ArgglyHtSasPhs 105
316 AAGCCAGCGCTCGGGGAGGCG 336
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 ArgHtSglYLySgLYcysgLY 112

```

seq\_name: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:US-09-492-026-7

seq\_documentation\_block:

; Sequence 7, Application US/09492026A

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; Bandman, Olga

; Corley, Neil C.

; Lal, Preeti

; Shah, Purvi

; TITLE OF INVENTION: HUMAN S100 PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/492.026A

; FILING DATE: 26-Jan-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Colette C. Muenzen

; REGISTRATION NUMBER: 39,784

; REFERENCE/DOCKET NUMBER: PF-0373 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-855-0555

; TELEFAX: 650-845-4166

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 113 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 488157

; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-492-026-7

alignment\_scores:

Quality: 355.50

Ratio: 3.864

Percent Similarity: 85.981

Percent Identity: 64.486

alignment\_block:

US-09-806-382a-2 x US-09-492-026-7

4

Align seq 1/1 to: US-09-492-026-7 from: 1 to: 113

```

16 TCCACAGCTGGAACGCAACATAGACACCATCATCAACACCTTCCACCAATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 SerGlnleuIdIuArgSerIleSerThrIleIleasnValPheHtSglntY 23
CCTGTGAGAGCTGGGACACCCAGACACCCGTGAGACCGAGGGGGAATTCGAAG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 tSerArgLyS TyrGlyHtSProaspHtIleuasnLysAlaGluPheHtSg 40
AGCTGTGCGAAGAAAGATCTGCAAAATTTTCTCAAGAGAGGAATTAAGAT 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 GAAAGGTATAGAACATCATGAGAGAGCTTGACACACAATTCACAA 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 GluAsnleuIuArgAspIleMetGluAspLeuAspThrAsnIuAspAs 73
216 GCAGCTGACCTTCGAGAGTTCATGCTGATGCGAGGCTAACCTGGG 265
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73 nglnleuserpneglucLucysmetleuMetglyLysleuIlephe 90
266 CCTCCACAGAGAAGATGACAGAGGTGACGAGGCCCTGGCCACACCAT 315
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90 lacyshtsluLysleuHtSgluasnAhnPro...ArgglyHtSasPhs 105
316 AAGCCAGCGCTCGGGGAGGCG 336
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 ArgHtSglYLySgLYcysgLY 112

```

seq\_name: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep:PCT-US01-00663-34972

seq\_documentation\_block:

; Sequence 34972, Application PC/TUS0100663

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

; FILE REFERENCE: PB 0004 WO 7

; CURRENT APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 04 February 2000 (04.02.00)

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 26 May 2000 (26.05.00)

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 03 August 2000 (03.08.00)

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 27 September 2000 (27.09.00)

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 21 September 2000 (21.09.00)

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 30 June 2000 (30.06.00)

; NUMBER OF SEQ ID NOS: 38837

; SOFTWARE: Molecular Dynamics Sequence Listing Engine

; SEQ ID NO 34972

; LENGTH: 64

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE: MAP TO AC011666.18

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2

; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALU 2.00e-34

; OTHER INFORMATION: EST\_HUMAN HIT: AT120432.1, EVALU 2.00e-33

PCT-US01-00663-34972

alignment\_scores:                      Length:                      64  
Quality:                      351.00                      Gaps:                      0  
Ratio:                      5.484  
Percent Similarity: 100.000                      Percent Identity: 100.000

## alignment\_block:

US-09-806-382a-2 x PCT-US01-00663-34972 ..

Align seg 1/1 to: PCT-US01-00663-34972 from: 1 to: 64

```
151 AAGGAGATTAAGATGAAGAGTGCATAGACATCATGTGAGAGCTGGA 200
|||||
1 LysGluAsnLysAsnGlnLysValIleGlnHisIleMetGluAspLeuAs 17
201 CCAATGTCAGACAGCAGCTGAGTTCGAGAGATTCAATCATGTCGATGG 250
|||||
17 phtAsnAlaAspLysGlnLeuSerPheGlnIuPheIleMetLeuMetA 34
251 CGAGGCTAACCTGGGCTCCAGCAGAGATGCACGAGGGTGAAGAGGC 300
|||||
34 laArgLeuThrTrpAlaSerHisGlnLysMetHisGlnLysAspGlnGly 50
301 CCTGGCCACCATTAAGCCAGGCTCGGGAGGAGCCACCCCC 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGlnGlyThrPro 64
```

seq\_name: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:US-09-864-761-40349

## seq\_documentation\_block:

```
/ Sequence 40349, Application US/09864761
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wenheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
/ FILE REFERENCE: Aecm1ca-X-1
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
```

```
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 40349
/ LENGTH: 64
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
```

```
/ OTHER INFORMATION: MAP TO AC011666.18
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1e+02
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.2
/ OTHER INFORMATION: SWISSPROT HIT: P06702, EVALUATE 2.00e-34
/ OTHER INFORMATION: EST_HUMAN HIT: A1720432.1, EVALUATE 2.00e-33
/ US-09-864-761-40349
```

alignment\_scores:                      Length:                      64  
Quality:                      351.00                      Gaps:                      0  
Ratio:                      5.484  
Percent Similarity: 100.000                      Percent Identity: 100.000

## alignment\_block:

US-09-806-382a-2 x US-09-864-761-40349 ..

Align seg 1/1 to: US-09-864-761-40349 from: 1 to: 64

```
151 AAGGAGATTAAGATGAAGAGTGCATAGACATCATGTGAGAGCTGGA 200
|||||
1 LysGluAsnLysAsnGlnLysValIleGlnHisIleMetGluAspLeuAs 17
201 CCAATGTCAGACAGCAGCTGAGTTCGAGAGATTCAATCATGTCGATGG 250
|||||
17 phtAsnAlaAspLysGlnLeuSerPheGlnIuPheIleMetLeuMetA 34
251 CGAGGCTAACCTGGGCTCCAGCAGAGATGCACGAGGGTGAAGAGGC 300
|||||
34 laArgLeuThrTrpAlaSerHisGlnLysMetHisGlnLysAspGlnGly 50
301 CCTGGCCACCATTAAGCCAGGCTCGGGAGGAGCCACCCCC 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGlnGlyThrPro 64
```

seq\_name: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:US-09-714-593-101

## seq\_documentation\_block:

```
/ Sequence 101, Application US/09714593
/ GENERAL INFORMATION:
/ APPLICANT: Wolf, Eckard
/ APPLICANT: Werner, Sabine
/ APPLICANT: Halle, Jorne-Peter
/ APPLICANT: Regenbogen, Johannes
/ APPLICANT: Goppelt, Andreas
/ TITLE OF INVENTION: Use of polypeptides or nucleic acids
/ TITLE OF INVENTION: encoding these for the diagnosis or treatment of skin
/ TITLE OF INVENTION: disorders, and their use for the identification of
/ TITLE OF INVENTION: pharmacologically active substances
/ FILE REFERENCE: 50125/009002
/ CURRENT FILING DATE: 2001-11-16
/ PRIOR APPLICATION NUMBER: US/09/714,593
/ PRIOR FILING DATE: 2001-11-16
/ PRIOR APPLICATION NUMBER: DE 19955349.1-41
/ PRIOR FILING DATE: 1999-11-17
/ PRIOR APPLICATION NUMBER: US 60/172,511
/ PRIOR FILING DATE: 1999-12-17
/ PRIOR APPLICATION NUMBER: DE 10030149.5-41
/ PRIOR FILING DATE: 2000-06-20
```

```

; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-714-593-101

```

```

alignment_scores:
  Quality: 329.50      Length: 107
  Ratio: 3.702         Gaps: 1
  Percent Similarity: 83.178  Percent Identity: 59.813

```

```

alignment_block:
US-09-806-382a-2 x US-09-714-593-101 ..

```

```

Align seg 1/1 to: US-09-714-593-101 from: 1 to: 113

```

```

16 TCGACGCTGGAACGCAACATAGACACATCATCAACACCTTCACCAATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 SerGImetGluArgSerIleThrIleIleAspThrPheHisGlnTyr 23
66 CTTGTGGAAGCTGGGGCACCACGACACCTTGACCAAGGGGGAATTCAAAG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 rSerHrGlySGluGlyHisProAspThrLeuSerTyrSGluPheArg 40
116 ACCTGCTGGAAGAAAGATCTGCAAAATTTCTCAGAGAGAGAGATAGAT 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
40 lmeValGlnIuIaGlnIleuAlaIrrPheMetLysLysGlnIuSarGAsn 56
166 GAAAAAGCTATGAAACACATCTGAGAGACCTGACACAAATGACAGACA 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 GluAlaIleuIleAsnAspIleMetGlnAspLeuAspThrAsnGlnAspAs 73
216 GCAGCTGAGCTTCGAGAGTTTCATGATGCTGATGCGGAGGCTACCTGG 265
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73 nGlnIleuSerPheGlnIuGlyCysMetMetLeuMetAlaLysLeuIlePhe 90
266 CCTCCGACGAGAAAGATGACAGAGGCTGAC...GAGGGCCCTGGGACCCAC 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90 lAcYSHISGlnIuLysLeuHisGlnAsnAsnProIrrGlyHisGlyHisSer 106
313 CATAGCCAGGCGCTGGGGAG 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
107 HisGlyLysGlyCysGlyLys 113

```

```

seq_name: /cgn2_6/prodata/2/paa/US60_COMB.pep:US-60-340-187-1029

```

```

seq_documentation_block:

```

```

; Sequence 1029, Application US/60340187
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Zhwei
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ghosh, Malalika
; APPLICANT: Meng, Gezhi
; APPLICANT: Boyle, Bryan J
; APPLICANT: Drmanac, Radolje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 813
; CURRENT APPLICATION NUMBER: US/60/340,187
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623

```

```

; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/03800
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/04927
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/04941
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/US01/08656
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/14827
; PRIOR FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 1192
; SOFTWARE: PL_FL_genes Version 6.0
; SEQ ID NO 1029
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(111)
; OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the table 8 as set
US-60-340-187-1029

```

```

alignment_scores:

```

```

  Quality: 320.00      Length: 93
  Ratio: 4.211         Gaps: 0
  Percent Similarity: 81.720  Percent Identity: 67.742

```

```

alignment_block:
US-09-806-382a-2 x US-60-340-187-1029 ..

```

```

Align seg 1/1 to: US-60-340-187-1029 from: 1 to: 111

```

```

64 TACTGTGAGCTGGGGCACCACGACACCTGGAACGAGGGGCAATTCAA 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
19 pheSerTyrLysLeuMet**AsnThrAspPhePheLysGlyArgArgGln 35
114 AGAGCTGTGCGAAGAAAGATCGCAAAATTTCTCAAGAGAGAGATAGGA 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
35 uGlnTyrLeuValGlnIleuLeuCysAsnLeuIleSerLysCysPhe**A 52
164 ATGAAAAGCTCATAGACACATCATGAGAGACCTGGACACAAATTCACAG 213
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
52 snAspThrThrGlyGlnHisIleMetGlnAspLeuAspThrAsnAlaAsp 68
214 AAGCAGCTGAGCTGAGAGAGTTTCATCATGATGAGGAGGAGGCTTAACCTG 263
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 LysGlnIleuSerPheGlnIuPheIleMetLeuMetAlaIrrGluThrTr 85

```

```

: LENGTH: 111
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(111)
: OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the table 8 as set
US-60-340-187-1030

alignment_scores:
    Quality: 320.00      Length: 93
    Ratio: 4.211        Gaps: 0
    Percent Similarity: 81.720      Percent Identity: 67.742

alignment_block:
US-09-806-382A-2 x US-60-340-187-1030      ..

Align seg 1/1 to: US-60-340-187-1030 from: 1 to: 111

64 TACTGTGGAAGCTGGGGCACCCAGACACCCCTGACACGGGGGAATTCAA 113
   ::::| |::::| |::: |::|::: |::
19 PheSerTyrLysLeuMet**AsnThrAspPhePheLysGlyArgArgG1 35
   ::::: |::: |::| |::|::: |::| |
114 AGAGCTGGTGGGAAAAGATCTCGCAAAATTTTCACAGACGAGAAATAGA 163
   ::::: |::: |::| |::|::: |::| |
35 GdnIlyrLeuValGlnLeuLeuCysAsnLeuLleSerLysCysPhe***A 52
   ::::: |::: |::| |::|::: |::| |
164 ATGAAGAAGGTCATAGAACATCATGTGAGAGACCTGACACAAATGACAG 213
   ||::: |::: |::| |::|::: |::| |
52 snAspThrThrGlyGlnHisIleMetGlnAspLeuAspThrAsnAlaAsp 68
   ||::: |::: |::| |::|::: |::| |
214 AAGCAGCTGAGGCTTCGAGAGGATTCATCATGCTGTATGGCGAGGCTAAC 263
   ||::: |::: |::| |::|::: |::| |
69 LysGlnLeuSerPheGlnGluPheIleMetLeuMetAlaArgLeuThrTr 85
   ||::: |::: |::| |::|::: |::| |
264 GGCCCTCCCAAGAGATGACAGAGGGTGAAGAGGCGCTCGGCACACACC 313
   ||::: |::: |::| |::|::: |::| |
85 PalaSerHisGlnLysMetHisGlnGlyAspGlnGlyProGlyHisHisH 102
   ||::: |::: |::| |::|::: |::| |
102 IsLysProGlyLeuGlnGlyGlnGlyThrPro 111
   ||::: |::: |::| |::|::: |::| |

seq_name: /cgn2_6/plotdata/2/paa/US60_COMB pep: US-60-340-187-1031

seq_documentation_block:
: Sequence 1031, Application US/60340187
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Wang, Zhiwei
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Ghosh, Malabika
: APPLICANT: Weng, Gezhi
: APPLICANT: Boyle, Bryan J
: APPLICANT: Drmanac, Radoje T
: TITLE OF INVENTION: Novel Nucleic Acids and
: FILE REFERENCE: 813
: CURRENT APPLICATION NUMBER: US/60/340,187
: CURRENT FILING DATE: 2001-12-12
: PRIOR APPLICATION NUMBER: US 09/488,725
: PRIOR FILING DATE: 2000-01-21
: PRIOR APPLICATION NUMBER: US 09/552,317
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: PCT/US00/35017
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/491,404
: PRIOR FILING DATE: 2000-01-25
: PRIOR APPLICATION NUMBER: PCT/US01/02623
: PRIOR FILING DATE: 2001-01-25
: PRIOR APPLICATION NUMBER: US 09/496,914
: PRIOR FILING DATE: 2000-02-03

```

```

? PRIOR APPLICATION NUMBER: US 09/560,875
? PRIOR FILING DATE: 2000-04-27
? PRIOR APPLICATION NUMBER: PCT/US01/03800
? PRIOR FILING DATE: 2001-02-05
? PRIOR APPLICATION NUMBER: US 09/515,126
? PRIOR FILING DATE: 2000-02-28
? PRIOR APPLICATION NUMBER: US 09/577,409
? PRIOR FILING DATE: 2000-05-18
? PRIOR APPLICATION NUMBER: PCT/US01/04927
? PRIOR FILING DATE: 2001-02-26
? PRIOR APPLICATION NUMBER: US 09/519,705
? PRIOR FILING DATE: 2000-03-07
? PRIOR APPLICATION NUMBER: US 09/574,454
? PRIOR FILING DATE: 2000-05-19
? PRIOR APPLICATION NUMBER: PCT/US01/04941
? PRIOR FILING DATE: 2001-03-05
? PRIOR APPLICATION NUMBER: US 09/540,217
? PRIOR FILING DATE: 2000-03-31
? PRIOR APPLICATION NUMBER: US 09/649,167
? PRIOR FILING DATE: 2000-08-23
? PRIOR APPLICATION NUMBER: PCT/US01/08631
? PRIOR FILING DATE: 2001-03-30
? PRIOR APPLICATION NUMBER: US 09/552,929
? PRIOR FILING DATE: 2000-04-18
? PRIOR APPLICATION NUMBER: US 09/770,160
? PRIOR FILING DATE: 2001-01-26
? PRIOR APPLICATION NUMBER: PCT/US01/08656
? PRIOR FILING DATE: 2001-04-18
? PRIOR APPLICATION NUMBER: US 09/577,408
? PRIOR FILING DATE: 2000-05-18
? PRIOR APPLICATION NUMBER: PCT/US01/14827
? PRIOR FILING DATE: 2001-05-16
? NUMBER OF SEQ ID NOS: 1192
? SOFTWARE: Pfam Version 6.0
? SEQ ID NO 1031
? LENGTH: 111
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(111)
? OTHER INFORMATION: Xaa = any amino acid or symbol as shown in the table 8 as set for
US-60-340-187-1031
```

```

alignment_scores:
    Quality: 320.00      Length: 93
    Ratio: 4.211        Gaps: 0
Percent Similarity: 81.720    Percent Identity: 67.742
```

alignment\_block:  
US-09-806-382a-2 x US-60-340-187-1031 ..

Align seg 1/1 to: US-60-340-187-1031 from: 1 to: 111

```

64 TACCTGTGGAAGCTGGGACCCAGACACCTGAAACCGAGGGGAATTCAA 113
   ::::: ||||| :::::
19 PheSerTyrLysLeuMet**AsnThrAspPhePheLysGlyArgArgG1 35
   ::::: ::::: ||| ||| :::::
114 AGAGCTGTGCGAAGAAGATCTGCAGAAATTTTCTCAGAGAAGAAATAGA 163
   ::::: ::::: ||| ||| :::::
35 uGlnTyrLeuValGlnLeuLeuCysAsnLeuIleSerLysCysPhe**A 52
   ::::: ::::: ||| ||| :::::
164 ATGAAGAAGTCATAGACACATCATGAGAGCCTGACACAAATGACAGAC 213
   ::::: ::::: ||||| ||||| :::::
52 snAspThrThrGlyGlnHisIleMetGluAspPheAsnAlaAsp 68
   ::::: ::::: ||||| ||||| :::::
214 AAGCAGCTGAGCTGAGAGAGTTATCTATGCTGATGGCGAGGCTAACCTG 263
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
69 LysGlnLeuSerPheGlnGluPheIleMetLeuMetAlaArgLeuThr 85
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 GGCCTCCACGAGAAAGATGACGAGGAGGTGACGAGGCGCCCTGGCCACACC 313
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

85 palAserHisGlnLysMetHisGlnGlyAspGlnGlyProGlyHisHisH 102
314 ATAGCCGAGGCTGGGAGGAGCCACCC 342
102 IsLysProGlyLeuGlnGlyGlnThrPro 111
```

seq\_name: /cgn2\_6/prodata/2/paa/PCTUS\_COMB.pep:PCT-US01-08656-6984

```

seq_documentation_block:
? Sequence 6984, Application PC/US0108656
? GENERAL INFORMATION:
? APPLICANT: HySeq, Inc
? TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
? FILE REFERENCE: 21272-066
? CURRENT APPLICATION NUMBER: PCT/US01/08656
? CURRENT FILING DATE: 2001-04-16
? PRIOR APPLICATION NUMBER: 09/522,929
? PRIOR FILING DATE: 2000-04-18
? PRIOR APPLICATION NUMBER: 09/770,160
? PRIOR FILING DATE: 2001-01-26
? NUMBER OF SEQ ID NOS: 10994
? SOFTWARE: Custom
? SEQ ID NO 6984
? LENGTH: 153
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: DOMAIN
? LOCATION: (38)..(75)
? OTHER INFORMATION: S-100/ICABP type calcium binding protein domain identified by
? OTHER INFORMATION: EMATRIX, accession number BL00303A, p-value=8.525e-12, raw sco
? OTHER INFORMATION: of 21.77
? NAME/KEY: DOMAIN
? LOCATION: (36)..(86)
? OTHER INFORMATION: S-100/ICABP type calcium binding domain identified by Pfam,
? OTHER INFORMATION: accession name S_100, E-value=0.0021, Pfam score of 21.6
? NAME/KEY: misc_feature
? LOCATION: (1)...(153)
? OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08656-6984
```

```

alignment_scores:
    Quality: 315.50      Length: 122
    Ratio: 3.286        Gaps: 7
Percent Similarity: 78.689    Percent Identity: 63.115
```

alignment\_block:  
US-09-806-382a-2 x PCT-US01-08656-6984 ..

Align seg 1/1 to: PCT-US01-08656-6984 from: 1 to: 153

```

1 ATGACTTGCAAAATGTGCGCAG...CTGGAACGCAACATGAGACCTCAT 47
   ||||| ||| ||| ::::: |||||
32 MetThrCysGlnMetProGlnHisValThrGlnGln**ArgProIle1 48
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 CAACACCTTCACCAATATCTGTGAAGCTGGGGACCCAGACACCTGA 97
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 eAsnThrSerHisGlnThrSerValIlyLysLeuGlnHisProAspThrLeuA 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 AAGAAGAGAAATAAGAAATGAAGGTCATGAAACATCATATGAGAGACT 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82 LysLysGlnAsnLysAsnGlnLysValIleGlnHisIleMetArgGlyPr 98
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 GGACACA...AATGCAGACAG.....CAGCTGAGCTGCGAGAGTTCA 238
   ||| ||||| ::::: ||||| ||||| ||||| ||||| ||||| |||||
98 ogLThrGlnAsnAlaIleGlnProAlaGlnLeuSer...ArgGluPheI 114
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 TCATGCTGATGGCGAGGCTAACCTGGGCTCC.....CAGGAGAAGATG 282
```









```
; SEQ ID NO 139803
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-139803
```

```
alignment_scores:
    Quality: 378.50      Length: 109
    Ratio: 4.027         Gaps: 1
    Percent Similarity: 86.239      Percent Identity: 69.725
```

alignment\_block:

US-09-806-382A-2 x US-09-791-537-139803 ..

Align seg 1/1 to: US-09-791-537-139803 from: 1 to: 122

```
13 ATGTGCGACGTGGACGACACATAGACCATCATCAACACCTTCACCA 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MetSerGlnMetGlnSerSerIleGlnThrIleIleAsnIlePheHisG1 17
63 ATACTCTGTGAAGCTGGGCGACCCACACACCCCTGAACGAGGGGAATTCA 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 ntyrSerValArgLeuGlyHisTyrAspThrLeuIleGlnLysGlnSerL 34
113 AAGACGTGCGCAAAAGATCTGCAAAATTTCTCAAGAAGAGAAATTAAG 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 ysgLInleuValGlnLysGlnLeuProAsnPheLeuLysGlnLysLys 50
163 AATGAAAGGTCAATAGACACATCATGAGAGACCTGAGACAAATGACAGA 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 AasnGlnAlaIleAsnGlnIleMetGlnAspLeuAspThrAsnValAs 67
213 CAAGCAGCTGAGCTTGAGAGATTTCATGCTGATGCGGAGGCTAACCT 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 ptyGlnLeuSerPheGlnGlnPheIleMetLeuValAlaArgLeuThrV 84
263 GGGCTCCCGACGAGAGATGACGAG...GGTGACGAGGCGCCCTGCGCAC 309
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84 aAlaSerHisGlnGlnMetHisAsnThrAlaProProGlnGlnLysHis 100
310 CACCATAGCCAGGCGCTCGGGGAGGCG 336
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 ArgHisGlnLysProGlnLysGlnLysGly 109
```

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-205-219-161

seq\_documentation\_block:

```
; Sequence 161, Application US/10205219
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alastair
; APPLICANT: Brooksbank, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205, 219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIORITY FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 161
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Intracellular calcium binding protein
US-10-205-219-161
```

alignment\_scores:

```
Quality: 355.50      Length: 107
Ratio: 3.864         Gaps: 1
Percent Similarity: 85.981      Percent Identity: 64.486
```

alignment\_block:

US-09-806-382A-2 x US-10-205-219-161 ..

Align seg 1/1 to: US-10-205-219-161 from: 1 to: 112

```
16 TCCGACCTGGACGACACATAGACCATCATCAACACCTTCACCAATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7 SerGlnLeuGlnArgSerIleSerThrIleIleAsnValPheHisGlnTy 23
66 CTCTGTGAAGCTGGGCGACCCACAGACCCCTGACGAGGGGGAATTCAAG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
23 rSerArgLysTyrGlnHisProAspThrLeuAsnLysAlaGlnPheLysG 40
116 ACCTGCTGCGCAAAAGATCTGCAAAATTTCTCAAGAAGAGAAATTAAGAT 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
40 lMetValAsnLysAspLeuProAsnPheLeuLysArgGlnLysArgAsn 56
166 GAAAGGTCAATAGACACATCATGAGAGACCTGACACAAATGACAGAA 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
57 GlnAsnLeuLeuArgAspIleMetGlnAspLeuAspThrAsnGlnAspAs 73
216 GCAGCTGAGCTTGAGAGATTTCATGCTGATGCGGAGGCTAACCTGGG 265
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
73 nGlnLeuSerPheGlnGlnLysGlnMetLeuMetGlyLysLeuIlePhe 90
266 CTTCCCGACGAGAGATGACGAGGAGGCGCCCTGCGCACACCAT 315
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
90 lacyHisGlnLysLysGlnLysGly 112
316 AAGCCAGGCTCGGGGAGGCG 336
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 ArgHisGlnLysGlnLysGly 112
```

seq\_name: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:US-09-791-537-139519

seq\_documentation\_block:

```
; Sequence 139519, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomedix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 139519
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-139519
```

```
alignment_scores:
    Quality: 355.50      Length: 107
    Ratio: 3.864         Gaps: 1
    Percent Similarity: 85.981      Percent Identity: 64.486
```

alignment\_block:

US-09-806-382A-2 x US-09-791-537-139519 ..

Align seg 1/1 to: US-09-791-537-139519 from: 1 to: 113

```
16 TCCGACCTGGACGACACATAGACCATCATCAACACCTTCACCAATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7 SerGlnLeuGlnArgSerIleSerThrIleIleAsnValPheHisGlnTy 23
```

```

66 CTCGTGAGCTGGGGCACCAGACACCTGAACGAGGGGAATTCAAG 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 rSerAglYsTYrGLYhISpTosPThLeuASnLYsAlaGLuPhLeYsG 40
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 AGCTGTGGGAAAAGATCTGCAAAATTTTCTCAGAAGGAGTAAGAAT 165
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
40 lUeTValSnLYsAsPLeuProASnPhLeuLYsArgGLuLYsATgaSn 56
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166 GAAAGGTCATAGACACATCATGAGACCTGGACACAAATGACAGACA 215
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
57 GluASnLeuAtrgAsPLeuMetGLuAsPLeuAsPThrASnGLuAsPAs 73
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 GCAGCTGAGGTTGAGGATCATGCTGATGGGAGGAGGTTAACTGGG 265
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 nGlnLeuSerPhLeuGLuLYsMetLeuMetGLuLYsLeu11ePha 90
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
266 CCTCCACAGAGAAGATGACAGAGGTGACAGAGGCCCTGCCACACCAT 315
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
90 lAcYshISGLuLYsLeuHISGLuASnASnPro...ArgGLYhISAsPHis 105
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
316 AGCCAGGCTCGGGGAGGCC 336
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
106 ArgHISGLuLYsGLuCYsGLY 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-182-995-26821
seq_documentation_block:
; Sequence 26821, Application US/10182995
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 1
; CURRENT APPLICATION NUMBER: US/10/182,995
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; NUMBER OF SEQ ID NOS: 29119
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 26821
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011666.18
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALUATION 2.00e-34
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: A1720432.1, EVALUATION 2.00e-33
; US-10-182-995-26821

```

```

alignment_scores:
Quality: 351.00      Length: 64
Ratio: 5.484        Gaps: 0

```

```

Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-09-806-382a-2 x US-10-182-995-26821
Align seg 1/1 to: US-10-182-995-26821 from: 1 to: 64
151 AAGGAGATTAAGATGAAGAAGTTCATAGACACATCATGAGAGACCTGGA 200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 LysGLuASnLYsASnGLuLYsVal11eGLuHIS11eMetGLuASnLeuAS 17
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 CACAATGCAGACAAGCAGCTGAGCTTCGAGAGTTCATGCTGATGG 250
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 PThrASnAlaSPrySGlnLeuSerPhLeuGLuPhLeuMetLeuMetA 34
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 lAaArgLeuThrTPraLaserHISGLuYsMetHISGLuGLYAsPGLuGLY 50
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 CCTGCCACACACATTAAGCCAGGCTCGGGGAGGGGACACCCCC 342
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 ProGLYhIShIShISLYsProGLYLeuGLuGLuThrPro 64
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

seq_name: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:US-10-203-134-34897
seq_documentation_block:
; Sequence 34897, Application US/10203134
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: PB 0004 WO 6
; CURRENT APPLICATION NUMBER: US/10/203,134
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; NUMBER OF SEQ ID NOS: 38628
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 34897
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011666.18
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1e+02
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALUATION 2.00e-34
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: A1720432.1, EVALUATION 2.00e-33
; US-10-203-134-34897

```

```

alignment_scores:
Quality: 351.00      Length: 64
Ratio: 5.484        Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

```

## alignment\_block:

US-09-806-382A-2 x US-10-203-134-34897

Align seg 1/1 to: US-10-203-134-34897 from: 1 to: 64

```
151 AAGGAGAAATAGAAATGAAAGTCAATAGAACACATCATGAGAGACCTGGA 200
|||||
1  LysGlnuSnlysaSnGlnlyuValIleGlnHnSlleMetGlnuAspLeuAs 17
201 CACAAATGCGACAGACAGCAGCTGAGCTTCGAGAGAGTTGATCATGCTGATGG 250
|||||
17 pPrHnAsnAlaAspLysGlnLeuSerPheGlnGluPheIleMetLeuMetA 34
251 CGAGGCTAACTGGGCTCCACAGAGAAATGCACGAGGGTGAAGAGGC 300
|||||
34 IaArGleuNrHTrPrAlaSerHnIsGlnLysMetHnSlGlnGlyAspGlnGly 50
301 CCTGGCCACCACCATTAAGCCAGGCTCGGGAGGAGGCCACCC 342
|||||
51 ProGlyHnShnSlHnSlHnSlHnSlHnSlHnSlHnSlHnSlHnSlHnSlHn 64
```

seq\_name: /cgn2.6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-203-136-35026

## seq\_documentation\_block:

; Sequence 35026, Application US/10203136

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: PB 0004 WO 3

; CURRENT APPLICATION NUMBER: US/10/203,136

; CURRENT FILING DATE: 2002-08-02

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 04 February 2000 (04.02.00)

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 26 May 2000 (26.05.00)

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 03 August 2000 (03.08.00)

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 27 September 2000 (27.09.00)

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 21 September 2000 (21.09.00)

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 30 June 2000 (30.06.00)

; NUMBER OF SEQ ID NOS: 38578

; SOFTWARE: Molecular Dynamics Sequence Listing Engine

; SEQ ID NO 35026

; LENGTH: 64

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC011666.18

; FEATURE:

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8

; FEATURE:

; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALUE 2.00e-34

; FEATURE:

; OTHER INFORMATION: EST\_HUMAN HIT: A1720432.1, EVALUE 2.00e-33

; US-10-203-136-35026

## alignment\_scores:

Quality: 351.00 Length: 64  
Ratio: 5.484 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-806-382A-2 x US-10-203-136-35026

Align seg 1/1 to: US-10-203-136-35026 from: 1 to: 64

```
151 AAGGAGAAATAGAAATGAAAGTCAATAGAACACATCATGAGAGACCTGGA 200
|||||
1  LysGlnuSnlysaSnGlnlyuValIleGlnHnSlleMetGlnuAspLeuAs 17
201 CACAAATGCGACAGACAGCAGCTGAGCTTCGAGAGAGTTGATCATGCTGATGG 250
|||||
17 pPrHnAsnAlaAspLysGlnLeuSerPheGlnGluPheIleMetLeuMetA 34
251 CGAGGCTAACTGGGCTCCACAGAGAAATGCACGAGGGTGAAGAGGC 300
|||||
34 IaArGleuNrHTrPrAlaSerHnIsGlnLysMetHnSlGlnGlyAspGlnGly 50
301 CCTGGCCACCACCATTAAGCCAGGCTCGGGAGGAGGCCACCC 342
|||||
51 ProGlyHnShnSlHnSlHnSlHnSlHnSlHnSlHnSlHnSlHnSlHnSlHn 64
```

seq\_name: /cgn2.6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-182-993-33897

## seq\_documentation\_block:

; Sequence 33897, Application US/10182993

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

; FILE REFERENCE: PB 0004 WO 2

; CURRENT APPLICATION NUMBER: US/10/182,993

; CURRENT FILING DATE: 2002-08-02

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 04 February 2000 (04.02.00)

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 26 May 2000 (26.05.00)

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 03 August 2000 (03.08.00)

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 27 September 2000 (27.09.00)

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 21 September 2000 (21.09.00)

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 30 June 2000 (30.06.00)

; NUMBER OF SEQ ID NOS: 37811

; SOFTWARE: Molecular Dynamics Sequence Listing Engine

; SEQ ID NO 33897

; LENGTH: 64

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC011666.18

; FEATURE:

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4

; FEATURE:

; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALUE 2.00e-34

; FEATURE:

; OTHER INFORMATION: EST\_HUMAN HIT: A1720432.1, EVALUE 2.00e-33

; US-10-182-993-33897

## alignment\_scores:

Quality: 351.00 Length: 64  
Ratio: 5.484 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-806-382A-2 x US-10-182-993-33897 ..

Align seg 1/1 to: US-10-182-993-33897 from: 1 to: 64

```
151 AAGGAGATAGATGAAAGGTCATAGACACATCATGAGAGACCTGGA 200
|||||
1 LysGluAsnLysAsnGluLysValIleGluHisIleMetGluAspLeuAs 17
201 CACAAATGCAGACAGACGCTGAGCTTCGAGAGATTCATCATGCTGATGG 250
|||||
17 pThrAsnAlaAspLysGlnLeuSerPheGluIuPheIleMetLeuMetA 34
251 CGAGGCTAACCTGGGCTCCACAGAGAAGATGCACGAGGTTGACGAGGC 300
|||||
34 laArgLeuThrTrpAlaSerHisGlnLysMetHisGlnGluLysAspGluGly 50
301 CCGGCCACACCATTAAGCCAGGCTCGGGAGGCGACCCCC 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGluGlyThrPro 64
```

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-182-997-25178

seq\_documentation\_block:

```
; Sequence 25178, Application US/10182997
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 10
; CURRENT APPLICATION NUMBER: US/10/182,997
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 26941
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 25178
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011666.18
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALU2 2.00e-34
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: A1720432.1, EVALU2 2.00e-33
; US-10-182-997-25178
```

alignment\_scores:                   Quality: 351.00                   Length: 64  
                                  Ratio: 5.484                   Gaps: 0  
Percent Similarity: 100.000       Percent Identity: 100.000

alignment\_block:

US-09-806-382A-2 x US-10-182-997-25178 ..

Align seg 1/1 to: US-10-182-997-25178 from: 1 to: 64

```
151 AAGGAGATAGATGAAAGGTCATAGACACATCATGAGAGACCTGGA 200
|||||
1 LysGluAsnLysAsnGluLysValIleGluHisIleMetGluAspLeuAs 17
201 CACAAATGCAGACAGACGCTGAGCTTCGAGAGATTCATCATGCTGATGG 250
|||||
17 pThrAsnAlaAspLysGlnLeuSerPheGluIuPheIleMetLeuMetA 34
251 CGAGGCTAACCTGGGCTCCACAGAGAAGATGCACGAGGTTGACGAGGC 300
|||||
34 laArgLeuThrTrpAlaSerHisGlnLysMetHisGlnGluLysAspGluGly 50
301 CCGGCCACACCATTAAGCCAGGCTCGGGAGGCGACCCCC 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGluGlyThrPro 64
```

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-203-135-34094

seq\_documentation\_block:

```
; Sequence 34094, Application US/10203135
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: PB 0004 WO 5
; CURRENT APPLICATION NUMBER: US/10/203,135
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37012
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 34094
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011666.18
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.2
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALU2 2.00e-34
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: A1720432.1, EVALU2 2.00e-33
; US-10-203-135-34094
```

alignment\_scores:                   Quality: 351.00                   Length: 64  
                                  Ratio: 5.484                   Gaps: 0  
Percent Similarity: 100.000       Percent Identity: 100.000

alignment\_block:

US-09-806-382A-2 x US-10-203-135-34094 ..

Align seg 1/1 to: US-10-203-135-34094 from: 1 to: 64

```
151 AAGGAGATATAGATGAAAAGTCATAGAACATCATGAGGAGCCTGGA 200
|||||
1 LysGlnAsnLysAsnGlnLysValIleGlnHisIleMetGlnAspLeuAs 17
201 CACAAATGCAGACAGACAGCTGAGCTTCGAGGAGTTTCATCATGCTGATGG 250
|||||
17 pThrAsnAlaAspLysGlnLeuSerPheGlnLupheIleMetLeuMeTa 34
251 CGAGGCTAACCTGGGCTCCACGAGAGATGCACGAGGTCACGAGGC 300
|||||
34 laArgLeuThrTrpAlaSerHisGlnLysMetHisGlnLysAspGlnGly 50
301 CCTGGCCACACCATTAAGCCAGGCTCGGGAGGACACCC 342
|||||
51 ProGlnHisHisLysProGlyLeuGlyGlnGlyThrPro 64
```

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-203-137-34972

```
seq_documentation_block:
; Sequence 34972, Application US/10203137
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: US/10/203,137
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 34972
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011666.18
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALU 2.00e-34
; OTHER INFORMATION: EST_HUMAN HIT: AI720432.1, EVALU 2.00e-33
; US-10-203-137-34972
```

alignment\_scores:  
Quality: 351.00 Length: 64  
Ratio: 5.484 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-806-382A-2 x US-10-203-137-34972 ..

Align seg 1/1 to: US-10-203-137-34972 from: 1 to: 64

```
151 AAGGAGATATAGATGAAAAGTCATAGAACATCATGAGGAGCCTGGA 200
|||||
1 LysGlnAsnLysAsnGlnLysValIleGlnHisIleMetGlnAspLeuAs 17
201 CACAAATGCAGACAGACAGCTGAGCTTCGAGGAGTTTCATCATGCTGATGG 250
|||||
17 pThrAsnAlaAspLysGlnLeuSerPheGlnLupheIleMetLeuMeTa 34
251 CGAGGCTAACCTGGGCTCCACGAGAGATGCACGAGGTCACGAGGC 300
|||||
34 laArgLeuThrTrpAlaSerHisGlnLysMetHisGlnLysAspGlnGly 50
301 CCTGGCCACACCATTAAGCCAGGCTCGGGAGGACACCC 342
|||||
51 ProGlnHisHisLysProGlyLeuGlyGlnGlyThrPro 64
```

seq\_name: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:US-10-203-139-33565

```
seq_documentation_block:
; Sequence 33565, Application US/10203139
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: PB 0004 WO 4
; CURRENT APPLICATION NUMBER: US/10/203,139
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37156
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 33565
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011666.18
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P06702, EVALU 2.00e-34
; OTHER INFORMATION: EST_HUMAN HIT: AI720432.1, EVALU 2.00e-33
; US-10-203-139-33565
```

alignment\_scores:  
Quality: 351.00 Length: 64  
Ratio: 5.484 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-806-382A-2 x US-10-203-139-33565 ..

Align seg 1/1 to: US-10-203-139-33565 from: 1 to: 64



```
151 AAGGAGATAGATGATAAGGTCATAGACATCATGAGAGACCTGGA 200
|||||
1 LysGluAsnLysAsnGluYsValIleGluHisIleMetGluAspLeuAs 17
201 CACAATGACAGACAGACAGCTGAGCTGAGAGATTTCATGCTGATGG 250
|||||
17 pNhrsmAlaAspLysGlnLeuSerPheGluGluPheIleMetLeuMet 34
251 CGAGGCTAACTGGGCTCCCAAGAGATGACAGAGGGTGACGAGGC 300
|||||
34 laArgLeuThrTrpAlaSerHisGluYsMetHisGluGluYsAspL 50
301 CTTGGCCACACCATAGACAGAGCTCGGGGAGGACCCCC 342
|||||
51 ProGlyHisHisHisLysProGlyLeuGlyGluGlyThrPro 64

seq_name: /cgn2_6/ptodata/1/paa/us09_NEW_COMB.pep:us-09-791-537-139518

seq_documentation_block:
; Sequence 139518, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 139518
; LENGTH: 113
; TYPE: PR1
; ORGANISM: Mus musculus
US-09-791-537-139518

alignment_scores:
Quality: 329.50 Length: 107
Ratio: 3.702 Gaps: 1
Percent Similarity: 83.178 Percent Identity: 59.813

alignment_block:
US-09-806-382a-2 x US-09-791-537-139518 ..
Align seg 1/1 to: US-09-791-537-139518 from: 1 to: 113

16 TCGAGCTGGAGCGACATAGAGACCATCATCAACACCTTCCACCATTA 65
|||||
7 SerGlnMetGluArgSerIleThrThrIleIleAspThrPheHisGlnTy 23
66 CTCTGTGAAGCTGGGGCAGCCAGACACCTGAGACAGGGGGAATTGCAAG 115
|||||
23 rSerArgLysGluGluYsHisProAspThrLeuSerLysGluPheArg 40
116 AGCTGTGCGAAGAAAGATCTGCAAAATTTTCTCAAGAGAGAAATAGAA 165
|||||
40 lMetValGluAlaGlnLeuAlaThrPheMetLysLysGluLysArgAsn 56
166 GAAAGGTCATAGACACATCATGAGAGACCTGAGACACAAATGACAGAA 215
|||||
57 GluAlaLeuIleAsnAspIleMetGluAspLeuAspThrAsnGlnAspAs 73
216 GCAGCTGAGCTTGCAGAGATTTCATCATGCTGATGCGAGGCTAACCTGG 265
|||||
73 nGlnLeuSerPheGlnGluGluYsMetMetLeuMetAlaLysLeuIlePhe 90
266 CTTCCACAGAGAAAGATGACAGAGGGTGAC...GAGGGCCCTGGCCACAC 312
|||||
90 lAcYsHisGlnLysLeuHisGlnAsnAsnProArgGlyHisGlyHisSer 106
313 CATAGCCAGGCTCGGGGAG 333
```

```
107 HisGlyLysGlyCysGlyLys 113

seq_name: /cgn2_6/ptodata/1/paa/us10_NEW_COMB.pep:us-10-134-841-3

seq_documentation_block:
; Sequence 3, Application US/10134841
; GENERAL INFORMATION:
; APPLICANT: HALLE, JORN-PETER
; APPLICANT: GOPELT, ANDREAS
; TITLE OF INVENTION: MRP8/MRP14 heterodimer, or its
; TITLE OF INVENTION: individual components in combination, for treating and/or
; TITLE OF INVENTION: preventing skin diseases, wounds and/or wound-healing
; TITLE OF INVENTION: disturbances, having a reduced quantity of MRP8/MRP14
; FILE REFERENCE: 50125/031002
; CURRENT APPLICATION NUMBER: US/10/134,841
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/322,925
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: DE 10121254.2
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 113
; TYPE: PR1
; ORGANISM: Mus musculus
US-10-134-841-3

alignment_scores:
Quality: 329.50 Length: 107
Ratio: 3.702 Gaps: 1
Percent Similarity: 83.178 Percent Identity: 59.813

alignment_block:
US-09-806-382a-2 x US-10-134-841-3 ..
Align seg 1/1 to: US-10-134-841-3 from: 1 to: 113

16 TCGAGCTGGAGCGACATAGAGACCATCATCAACACCTTCCACCATTA 65
|||||
7 SerGlnMetGluArgSerIleThrThrIleIleAspThrPheHisGlnTy 23
66 CTCTGTGAAGCTGGGGCAGCCAGACACCTGAGACAGGGGGAATTGCAAG 115
|||||
23 rSerArgLysGluGluYsHisProAspThrLeuSerLysGluPheArg 40
116 AGCTGTGCGAAGAAAGATCTGCAAAATTTTCTCAAGAGAGAAATAGAA 165
|||||
40 lMetValGluAlaGlnLeuAlaThrPheMetLysLysGluLysArgAsn 56
166 GAAAGGTCATAGACACATCATGAGAGACCTGAGACACAAATGACAGAA 215
|||||
57 GluAlaLeuIleAsnAspIleMetGluAspLeuAspThrAsnGlnAspAs 73
216 GCAGCTGAGCTTGCAGAGATTTCATCATGCTGATGCGAGGCTAACCTGG 265
|||||
73 nGlnLeuSerPheGlnGluGluYsMetMetLeuMetAlaLysLeuIlePhe 90
266 CTTCCACAGAGAAAGATGACAGAGGGTGAC...GAGGGCCCTGGCCACAC 312
|||||
90 lAcYsHisGlnLysLeuHisGlnAsnAsnProArgGlyHisGlyHisSer 106
313 CATAGCCAGGCTCGGGGAG 333
107 HisGlyLysGlyCysGlyLys 113

seq_name: /cgn2_6/ptodata/1/paa/us09_NEW_COMB.pep:us-09-791-537-13822
seq_documentation_block:
; Sequence 13822, Application US/09791537
```

```
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13822
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-791-537-13822
```

```
alignment_scores:
  Quality: 328.00      Length: 102
  Ratio: 3.859        Gaps: 1
  Percent Similarity: 83.333      Percent Identity: 63.725
```

alignment\_block:

US-09-806-382a-2 x US-09-791-537-13822 ..

Align seg 1/1 to: US-09-791-537-13822 from: 1 to: 118

```
43 ATCATCAACCTTCACCAATACCTGTGAAGCTGGGACCCAGACAC 92
|||||  |||||||  |||||||  |||||||  |||
1 IleIleAsnIlePheHisGlnTyrSerValArgValGlyProArgAspSe 17
93 CCTGAACCGAGGGGAATTCAAGAGCTGTGCGAATAAGATGTCAAAATT 142
:|||||  |||||||  |||||||  |||||||  |||||||  |||
17 rleuserGlnLysGlnPheLysGlnLeuValGlnLysGlnLeuHisAsn 34
143 TTCTCAAGAAGAGATAAGATGAAGATCAATAGACACATCATGAG 192
|||||  |||||||  |||||||  |||||||  |||
34 helLeuLysLysGlnAlaArgAspGlnLysAlaIleAsnAspIleMetGlu 50
193 GACCTGGAGACAAATGCAGACAGCTGAGCTTGAGAGAGTTCATCAT 242
|||||  |||||||  |||||||  |||||||  |||||||  |||
51 AspLeuAspThrAsnGlnAspLysGlnLeuSerPheGlnGluPheValI 67
243 GCTGATGGCGAGGCTAACCTGGGCTCCACGAGAGATGACAGAGGT. 291
:|||||  |||||||  |||||||  |||||||  |||||||  |||
67 eleuMetAlaArgLeuValHisAlaSerHisGlnGluMetHisLysAsn 84
292 .....GACGAGGGCCCTGGGCCACACCATTAAGCCAGGCTCGGGGAGG 336
:::  |||  |||  |||  |||  |||  |||  |||  |||  |||
84 laproHisAspHisGlnGlnHisSerHisGlyProGlnLeuGlnGlySer 100
337 ACCCC 342
|||
101 GlyPro 102
```

OM of: US-09-806-382a-2 to: Issued\_Patents\_AA:\* out-format : pfs  
Date: Sep 9, 2002 3:11 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frame+np.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09806382/runtat\_09092002.143848.15391/app-query.fasta.1.752  
-DB=Issued\_Patents\_AA -OFMT=fastaa -SUFFIX=ra1 -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEX=7.000 -START=1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09806382.ecgn1\_1\_39 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPHY -WAIT -THREADS=1

## Search information block:

Query: US-09-806-382a-2

Query length: 345

Database: Issued\_Patents\_AA.\*

Database sequences: 231628

Database length: 2442594

Search time (sec): 38.230000

## Score list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation	..	1.9e-33	113	1
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-918-727-7		355.50	733.90	1.9e-33	113	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-09-205-680A-7		355.50	733.90	1.9e-33	113	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-08-200-016-5		351.00	730.23	5.4e-33	64	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-200-016-5		262.00	546.64	1.2e-22	50	1			
/cgn2_6/ptodata/2/iaa/6B.COMB.pcp:US-08-568-910D-20		214.50	441.17	4.8e-17	92	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-568-910D-20		214.50	441.17	4.8e-17	92	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-568-910D-19		211.50	439.19	6.2e-17	91	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-08-568-910D-19		211.50	439.19	6.2e-17	92	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-568-910D-19		211.50	434.90	1.1e-16	92	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-08-568-910D-19		211.50	434.90	1.1e-16	92	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-918-727-5		177.50	363.82	9.7e-13	92	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-918-727-5		177.50	363.82	9.7e-13	92	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-09-051-589-1		173.50	355.45	2.8e-12	92	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-09-051-589-1		172.50	353.47	3.7e-12	91	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-468-946-2		150.50	306.42	1.4e-09	101	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-08-468-946-2		150.50	306.42	1.4e-09	101	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-468-946-2		150.50	306.42	1.4e-09	101	1			
/cgn2_6/ptodata/2/iaa/6B.COMB.pcp:US-08-468-946-2		149.50	303.94	1.8e-09	105	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-468-946-2		149.50	303.94	1.8e-09	105	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-07-987-272A-1		146.50	299.45	3.9e-09	88	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-07-987-272A-1		146.50	299.45	3.9e-09	89	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-918-727-3		138.00	280.59	3.9e-08	98	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-09-205-680A-3		138.00	280.59	3.9e-08	98	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-09-048-889-11		137.00	281.89	3.5e-08	26	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-09-048-889-11		137.00	281.89	3.5e-08	26	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-488-379-144		137.00	291.89	3.5e-08	26	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-488-379-144		137.00	291.89	3.5e-08	26	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-07-987-272A-16		133.50	271.71	1.3e-07	93	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-07-987-272A-16		133.50	271.71	1.3e-07	93	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-385-241-1		133.50	271.71	1.3e-07	93	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-385-241-1		133.50	271.71	1.3e-07	93	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-08-480-190-145		128.00	273.88	3.8e-07	24	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-480-190-145		128.00	273.88	3.8e-07	24	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-480-190-145		128.00	273.88	3.8e-07	24	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-480-190-145		128.00	273.88	3.8e-07	24	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-07-987-272A-17		126.50	269.11	8.0e-07	76	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-07-987-272A-17		126.50	269.11	8.0e-07	76	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-08-568-310D-2		126.00	262.09	8.1e-07	51	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-568-310D-2		126.00	262.09	8.1e-07	51	1			
/cgn2_6/ptodata/2/iaa/5A.COMB.pcp:US-07-987-272A-9		123.00	249.54	2.2e-06	95	1			
/cgn2_6/ptodata/2/iaa/5B.COMB.pcp:US-07-987-272A-9		123.00	249.54	2.2e-06	95	1			
/cgn2_6/ptodata/2/iaa/6A.COMB.pcp:US-08-918-727-1		117.00	236.18	1.1e-05	103	1			

/cgn2\_6/ptodata/2/iaa/6A.COMB.pcp:US-09-205-680A-1 + 117.00 236.18 1.1e-05 103  
/cgn2\_6/ptodata/2/iaa/5A.COMB.pcp:US-07-987-272A-12 + 113.00 231.02 3.0e-05 75  
/cgn2\_6/ptodata/2/iaa/5A.COMB.pcp:US-08-469-486-58 + 113.00 228.02 3.2e-05 101

seq\_name: /cgn2\_6/ptodata/2/iaa/5B.COMB.pcp:US-08-918-727-7

## seq\_documentation\_block:

Sequence 7, Application US/08918727  
Patent No. 5849528  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Corley, Neil C.  
APPLICANT: Lal, Preeti  
APPLICANT: Shah, Puri  
TITLE OF INVENTION: HUMAN S100 PROTEINS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/918,727  
FILING DATE: Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0373 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:

## INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 113 amino acids  
TYPE: amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 488157  
US-08-918-727-7

## alignment\_scores:

Quality: 355.50 Length: 107  
Ratio: 3.864 Gaps: 1  
Percent Similarity: 85.981 Percent Identity: 64.486

## alignment\_block:

US-09-806-382A-2 x US-08-918-727-7 ..

Align seg 1/1 to: US-08-918-727-7 from: 1 to: 113

16 TCGACGCTGGAAACGACATAGACATCATTAACACCTTCCACCAATA 65  
|||||  
7 SerGlnLeuGluArgSerIleSerThrIleIleAsnValPheHisGlnTy 23  
66 CTCCTGAGCGGGGCGGCGGACCGGACCGCTGACGAGGGGGAATTCAAG 115  
|||||  
23 TserArgIlystyIrgIlyHisProAspThrLeuAsnIlyValIacIuPheIlysg 40

```

116 AGCTGTGCGAAAGATCTGCAAAATTTTCTCAAGAAGAGATTAAGAT 165
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
40 lmetvalasnlvsaspleuproasnphenleuysarglulysargasn 56
166 GAAAGGTCATAGACACATCATGAGAGCCTGAGACACAATGACAGCAA 215
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
57 Gluasnlleuuarargaspllemetgluaspleuasprthrasnclnaspas 73
216 GCAGCTGAGCTGAGAGGATTCATGCTGATGGGAGGAGCTAAGCTGGG 265
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
73 nglnleuserpnegluglucysmetleuemetglulysleuilephea 90
266 CCTCCACGACGAGAGATGACAGAGGTGACGAGGCGCTGCGCACACCAT 315
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
90 lacyshtsglulysleuhtsglulasnasnpro...Argglyhtsasphts 105
316 AACCCAGCCCTCGGGAGGCG 336
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
106 Arghtsgllylsgllycysgly 112

```

seq\_name: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:US-09-205-680A-7

```

seq_documentation_block:
: Sequence 7, Application US/09205680A
: Patent No. 6103497
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Bandman, Olga
: APPLICANT: Corley, Neil C.
: APPLICANT: Lal, Preeti
: APPLICANT: Shah, Purvi
: TITLE OF INVENTION: HUMAN S100 PROTEINS
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/205,680A
: FILING DATE: Herewith
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Colette C. Muenzen
: REGISTRATION NUMBER: 39,784
: REFERENCE/DOCKET NUMBER: PF-0373 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 113 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: Genbank
: CLONE: 488157
: US-09-205-680A-7

```

alignment\_scores: Quality: 355.50 Length: 107  
Ratio: 3.864 Gaps: 1  
Percent Similarity: 85.981 Percent Identity: 64.486

alignment\_block:

US-09-806-382A-2 x US-09-205-680A-7 ..

Align seq 1/1 to: US-09-205-680A-7 from: 1 to: 113

```

16 TCGCAGCTGAGACGACACATAGACCATCATGACACCTTCACACCAT 65
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7 Serglleucluargerislerthrlelleasvalphthtsglntly 23
66 CTCTGTGAGAGCTGGGCGACCCACACACCCCTGAGACAGGGGATTTCAAG 115
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
23 rserarglvslytrglyhtsiproasprthreasnlyslalgluphts 40
116 AGCTGTGCGAAAGATCTGCAAAATTTTCTCAAGAAGAGATTAAGAT 165
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
40 lmetvalasnlvsaspleuproasnphenleuysarglulysargasn 56
166 GAAAGGTCATAGACACATCATGAGAGCCTGAGACACAATGACAGCAA 215
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
57 Gluasnlleuuarargaspllemetgluaspleuasprthrasnclnaspas 73
216 GCAGCTGAGCTGAGAGGATTCATGCTGATGGGAGGAGCTAAGCTGGG 265
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
73 nglnleuserpnegluglucysmetleuemetglulysleuilephea 90
266 CTTCCACGACGAGAGATGACAGAGGTGACGAGGCGCTGCGCACACCAT 315
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
90 lacyshtsglulysleuhtsglulasnasnpro...Argglyhtsasphts 105
316 AACCCAGCCCTCGGGAGGCG 336
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
106 Arghtsgllylsgllycysgly 112

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-200-016-6

```

seq_documentation_block:
: Sequence 6, Application US/08200016
: Patent No. 5614397
: GENERAL INFORMATION:
: APPLICANT: Weisman, Irving
: APPLICANT: Lagasse, Eric
: TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR MODULATING
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: STREET: 2200 Sand Hill Road, Suite 100
: CITY: Menlo Park
: STATE: California
: COUNTRY: US
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/200,016
: FILING DATE: 22-FEB-1994
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Rae-Venter, Barbara
: REGISTRATION NUMBER: 32,750
: REFERENCE/DOCKET NUMBER: 06037/003001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 854-5277
: TELEFAX: (415) 854-0875
: INFORMATION FOR SEQ ID NO: 6:

```

SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-200-016-6

alignment\_scores:  
Quality: 351.00 Length: 64  
Ratio: 5.484 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-806-382A-2 x US-08-200-016-6

Align seg 1/1 to: US-08-200-016-6 from: 1 to: 64

151 AAGGAGATATGATGAAAAAGTGCATAGAACACATCATGAGACCTGGA 200  
|||||  
1 LysGluAsnLysAsnGluValIleGluHisIleMetGluAspLeuAs 17  
201 CACAATGACACAGCAGCTGAGCTTCGAGAGTGCATCATGCTGATGG 250  
|||||  
17 pTrpAsnAlaAspLysGlnLeuSerPheGluGluPheIleMetLeuMet 34  
251 CGAGGCTAACCTGGGCTCCACGAGAGATGCAGAGGGTGACGAGGCG 300  
|||||  
34 IaArgLeuThrTrpAlaSerHisGluLysMetHisGluLysAspGluGly 50  
301 CCGGCGCACCCACCATTAAGCCAGGCTCGGGAGGCAACCC 342  
|||||  
51 ProGlyHisHisHisLysProGlyLeuGlyGluGlyThrPro 64

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-200-016-5

seq\_documentation\_block:

Sequence 5, Application US/08200016  
Patent No. 5614397  
GENERAL INFORMATION:  
APPLICANT: Weissman, Irving  
APPLICANT: Lagasse, Eric  
TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR MODULATING  
TITLE OF INVENTION: APOPTOSIS IN HEMATOLOGY CELLS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: California  
COUNTRY: US  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/200,016  
FILING DATE: 22-FEB-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Rae-Ventler, Barbara  
REGISTRATION NUMBER: 32,750  
REFERENCE/DOCKET NUMBER: 06037/003001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 854-5277  
TELEFAX: (415) 854-0875  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-200-016-5

alignment\_scores:  
Quality: 262.00 Length: 50  
Ratio: 5.240 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-806-382A-2 x US-08-200-016-5

Align seg 1/1 to: US-08-200-016-5 from: 1 to: 50

1 ATGACTTGCAAAATGTGCGACGTGGAGCAACATAGACCATCATCA 50  
|||||  
1 MethrCysLysMetSerGlnLeuGlnArgAsnIleGluThrIleIleAs 17  
17 nThrPheHisGlnTrpSerValLysLeuGlyHisProAspThrLeuAsn 34  
101 AGGGGATTTCAAGAGCTGTCGCAAAAGATCTGCATAATTTCTCAAG 150  
|||||  
34 InGlyGluPheLysGlnLeuValArgLysAspLeuGlnAsnPheLeuLys 50

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-568-310D-20

seq\_documentation\_block:

Sequence 20, Application US/08568310D  
Patent No. 5976832  
GENERAL INFORMATION:  
APPLICANT: HITOMI, JIRO  
APPLICANT: YAMAGUCHI, KEN  
APPLICANT: YAMAMURA, TOKUJIRO  
APPLICANT: KIMURA, TATSUJI  
TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE  
STREET: 99 PARK AVENUE  
CITY: NEW YORK CITY  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB  
MEDIUM TYPE: STORAGE  
COMPUTER: IBM-PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS 6.2  
SOFTWARE: WORDPERFECT 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/568,310D  
FILING DATE: DECEMBER 6, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 7-70468 and 7-45564 (both Japan)  
FILING DATE: 3/6/95 and 3/6/95, respectively  
ATTORNEY/AGENT INFORMATION:  
NAME: KLEIN, MILTON  
REGISTRATION NUMBER: 27101  
REFERENCE/DOCKET NUMBER: 3316  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)953-3350  
TELEFAX: (212)953-3352  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 92  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

PUBLICATION INFORMATION:  
RELEVANT RESIDUES IN SEQ ID NO: 20:  
US-08-568-310D-20 FROM 1 TO 92

alignment\_scores:  
Quality: 214.50 Length: 92  
Ratio: 2.979 Gaps: 1  
Percent Similarity: 78.261 Percent Identity: 46.739

alignment\_block:

US-09-806-382A-2 x US-08-568-310D-20 ..

Align seg 1/1 to: US-08-568-310D-20 from: 1 to: 92

```
13 ATGTCCAGCTGGAACGCAATAGACCATCATCAACACCTTCCACCA 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MetThrLysLeuGluGlnHisLeuGluGlyIleValAsnIlePheHsgl 17
63 ATACTGTGAAGCTGGGCGACCCAGACACCCGTGACAGGAGGAATTCA 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 nTyrSerValArgLysGlyHisPheAspThrLeuSerLysGlyGluLeu 34
113 AAGAGCTGTGCGAAAGATCTGCAAAATTTCTCAAGAAGAGAAATAG 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 ysgInleuLeuThrLysGluLeuAlaAsnThrIle..LysAsnIleLys 49
163 AATGAAAGGTGATAGACATCATGAGGAGCCTGACACCAATATCAGA 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
50 AspLysAlaValIleAspGluIlePheGlnGlyLeuAspAlaAsnGln 66
213 CAAGAGCTGAGCTTCGAGAGATTTCATCATGCTGATGGCGAGCTAAC 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 pgiuGlnValAspPheGlnGluPheIleSerLeuValAlaIleAlaLeu 83
263 GGGCCTCCACGAGAAAGATGCACGAG 288
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
83 ysaIaIaIaHisTyrHisThrHisLys 91
```

seq\_name: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:US-09-270-455-20

seq\_documentation\_block:

Sequence 20, Application US/09270455  
Patent No. 6313267

GENERAL INFORMATION:

APPLICANT: HITOMI, JIRO

APPLICANT: YAMAGUCHI, KEN

APPLICANT: YAMAMURA, TOKUJIRO

APPLICANT: KIMURA, TATSUJI

TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESS: 99 PARK AVENUE

STREET: 6th FLOOR

CITY: NEW YORK CITY

STATE: NEW YORK

COUNTRY: USA

ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB

COMPUTER: IBM-PC COMPATIBLE

OPERATING SYSTEM: PC-DOS 6.2

SOFTWARE: WORDPERFECT 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/270,455

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/568,310

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: KLEIN, MILTON

REGISTRATION NUMBER: 27101

REFERENCE/DOCKET NUMBER: 3316

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)953-3350

TELEFAX: (212)953-3352

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 92

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: CDNA

PUBLICATION INFORMATION:

RELEVANT RESIDUES IN SEQ

US-09-270-455-20

alignment\_scores:

Quality: 214.50 Length: 92  
Ratio: 2.979 Gaps: 1  
Percent Similarity: 78.261 Percent Identity: 46.739

alignment\_block:

US-09-806-382A-2 x US-09-270-455-20 ..

Align seg 1/1 to: US-09-270-455-20 from: 1 to: 92

```
13 ATGTCCAGCTGGAACGCAATAGACCATCATCAACACCTTCCACCA 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MetThrLysLeuGluGlnHisLeuGluGlyIleValAsnIlePheHsgl 17
63 ATACTGTGAAGCTGGGCGACCCAGACACCCGTGACAGGAGGAATTCA 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 nTyrSerValArgLysGlyHisPheAspThrLeuSerLysGlyGluLeu 34
113 AAGAGCTGTGCGAAAGATCTGCAAAATTTCTCAAGAAGAGAAATAG 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 ysgInleuLeuThrLysGluLeuAlaAsnThrIle..LysAsnIleLys 49
163 AATGAAAGGTGATAGACATCATGAGGAGCCTGACACCAATATCAGA 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
50 AspLysAlaValIleAspGluIlePheGlnGlyLeuAspAlaAsnGln 66
213 CAAGAGCTGAGCTTCGAGAGATTTCATCATGCTGATGGCGAGCTAAC 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 pgiuGlnValAspPheGlnGluPheIleSerLeuValAlaIleAlaLeu 83
263 GGGCCTCCACGAGAAAGATGCACGAG 288
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
83 ysaIaIaIaHisTyrHisThrHisLys 91
```

seq\_name: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:US-08-794-000-2

seq\_documentation\_block:

Sequence 2, Application US/08794000  
Patent No. 6087123

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Metal-Containing Ribonucleotide Polypeptides

NUMBER OF SEQUENCES: 4

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/794,000

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/DE96/01337

FILING DATE: 17-JUL-1996

```

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 25 992.0
; FILING DATE: 17-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 30 500.0
; FILING DATE: 18-AUG-1995
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-794-000-2

alignment_scores:
    Quality: 213.50      Length: 91
    Ratio: 2.847         Gaps: 1
    Percent Similarity: 82.418    Percent Identity: 45.055

alignment_block:
US-09-806-382A-2 x US-08-794-000-2 ..

Align seg 1/1 to: US-08-794-000-2 from: 1 to: 91

16 TCGAGCTGGAGCGCAACATGAGACCATCATCACACCTTCACCAATA 65
   ::::::::::::::::::::| | | | | | | | | | | | | | | |
1 ThrLysLeuGlnAspHisLeuGlnGlyLeuIleIleAsnIlePheHisGlnTy 17
66 CTCTGTGAAGCTGGGGCACCACCGACACCTGAAACAGGGGGGAATTCAAG 115
   || | | | | | | | | | | | | | | | | | | | | | | | |
17 rSerValArgLeuGlnGlyHisTyrAspThrLeuIleLysArgGlnLeuLysG 34
116 AGCTGGTGGCAAAAAGATCTGCAAAATTTCTCAAGAAAGAGGAATAAAT 165
   :: | | | | | | | | | | | | | | | | | | | | | | | |
34 InLeuIleThrLysGlnLeuProAsnThrLeu...LysAsnThrLysAsp 49
166 GAAAGGCTCATAGAACATCATGAGAGACCTGACACAAATGCAGACAA 215
   :: | | | | | | | | | | | | | | | | | | | | | | | |
50 GlnGlyThrIleAspLysIlePheGlnAsnLeuAspLalaAsnGlnAspGl 66
216 GCAGCTGAGCTTCGAGAGTTTCATCATGCTGATGGCGAGGCTTAACCTG 265
   : | | | | | | | | | | | | | | | | | | | | | | | |
66 uGlnValSerPheLysGlnPheValValLeuValThrAspValLeuLett 83
266 CCTCCACGAGAGATGACGACGAG 288
83 hrlAlaHisAspAsnIleHisLys 90

seq_name: /cgn2_6/plodata/2/1aa/5B_COMB.pep:US-08-568-310D-19

seq_documentation_block:
; Sequence 19, Application US/08568310D
; Patent No. 5976832
; GENERAL INFORMATION:
; APPLICANT: HITOMI, JIRO
; APPLICANT: YAMAGUCHI, KEN
; APPLICANT: YAMAMURA, TOKUJIRO
; APPLICANT: KIMURA, TATSUJI
; TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE
; STREET: 99 PARK AVENUE
; CITY: NEW YORK CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB
; MEDIUM TYPE: STORAGE
```

```

;
; COMPUTER: IBM-PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS 6.2
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,310D
; FILING DATE: DECEMBER 6, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7-70468 and 7-45564(both Japan)
; FILING DATE: 3/6/95 and 3/6/95, respectively
; ATTORNEY/AGENT INFORMATION:
; NAME: KLEIN, MILTON
; REGISTRATION NUMBER: 27101
; REFERENCE/DOCKET NUMBER: 3316
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)953-3350
; TELEFAX: (212)953-3352
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 19:
; FROM 1 TO 92
; US-08-568-310D-19

alignment_scores:
    Quality: 211.50      Length: 92
    Ratio: 2.858         Gaps: 1
    Percent Similarity: 80.435    Percent Identity: 44.565

alignment_block:
US-09-806-382A-2 x US-08-568-310D-19 ..

Align seg 1/1 to: US-08-568-310D-19 from: 1 to: 92

13 ANGTCGACCTGGAACGCAACATAGACCATCATCAACACCTTCACCA 62
   || | | | | | | | | | | | | | | | | | | | | | | | |
1 MetThrLysLeuGlnAspHisLeuGlnGlyLeuIleIleAsnIlePheHisGl 17
63 ATACTGTGGAAGCTGGGGCACCACCGACACCTGAAACAGGGGGGAATTCA 112
   || | | | | | | | | | | | | | | | | | | | | | | | |
17 nTyrSerValArgValGlnHisPheAspThrLeuAsnLysArgGlnLeuL 34
113 AAGAGCTGTGGCAAAAGATCTGCAAAATTTCTCAAGAAAGAGGAATAAG 162
   || | | | | | | | | | | | | | | | | | | | | | | | |
34 ySgInLeuIleThrLysGlnLeuProLysThrLeu...GlnAsnThrLys 49
213 AATGAAAAGCTCATAGAACATCATGAGAGACCTGAGACAAATGCAGAGA 212
   :: | | | | | | | | | | | | | | | | | | | | | | | |
50 AspGlnProThrIleAspLysIlePheGlnAspLeuAspLalaAspLysAs 66
213 CAAGCAGCTGAGCTTCGAGAGTTTCATCATGCTGATGGCGAGGCTAACCT 262
   : | | | | | | | | | | | | | | | | | | | | | | | |
66 pGlyAlaValSerPheGlnGlnPheValValLeuValSerArgValLeuL 83
263 GGGCTCCGACGAGAGATGACGACGAG 288
83 ySthrAlaHisIleAspIleHisLys 91

seq_name: /cgn2_6/plodata/2/1aa/6B_COMB.pep:US-09-270-455-19

seq_documentation_block:
; Sequence 19, Application US/09270455
; Patent No. 6313267
; GENERAL INFORMATION:
; APPLICANT: HITOMI, JIRO
; APPLICANT: YAMAGUCHI, KEN
; APPLICANT: YAMAMURA, TOKUJIRO
```

```
APPLICANT: KIMURA, TATSUJI
TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: WYATT, GERBER, MELLER & O'ROURKE
STREET: 99 PARK AVENUE
CITY: NEW YORK CITY
STATE: NEW YORK
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 Kb
MEDIUM TYPE: STORAGE
COMPUTER: IBM-PC COMPATIBLE
OPERATING SYSTEM: PC-DOS 6.2
SOFTWARE: WORDPERFECT 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/270,455
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/568,310
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KLEIN, MILTON
REGISTRATION NUMBER: 27101
REFERENCE/DOCKET NUMBER: 3316
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)953-3350
TELEFAX: (212)953-3352
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 92
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: CDNA
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 19: FROM 1 TO 92
US-09-270-455-19
```

```
alignment_scores:
Quality: 211.50      Length: 92
Ratio: 2.858         Gaps: 1
Percent Similarity: 80.435      Percent Identity: 44.565
```

alignment\_block:

US-09-806-382a-2 x US-09-270-455-19 ..

Align seg 1/1 to: US-09-270-455-19 from: 1 to: 92

```
13 ATGTGCGAGCTGGAACGACATAGAGACCATCATCAACACTTCCACGA 62
|||||
1 MetThrLysLeuGluAspHisLeuGluGlyIleIleHisnIlePheHisgl 17
63 ATACTCTGAAGCTGGGGGCAACCCAGACCCCTGAACCGAGGGAATTCA 112
|||||
17 nTyrSerValArgValGlyHisPheAspThrIleuAsnLysArgGluLeu 34
113 AAGAGCTGTCGCGAAAGATCTGCAAAATTTCTCAAGAAGAGAGANTAG 162
|||||
34 ysgInleuIleThrLysGluLeuProLysThrLeu...GlnAsnThrLys 49
163 AATGAAAGGCTCATAGAACACATCATGAGACCTGGACACAAATCCAGA 212
|||||
50 AspGlnProThrIleAspLysIlePheGlnAspLeuAspAlaAspLysAs 66
213 CAAGCAGCTAGCTTGAGAGATTCAATCATGCTGATGGCGAGGCTAACCT 262
|||||
66 pGlyAlaValSerPheGluIlePheValValLeuValSerArgValLeu 83
```

```
263 GGGCCTCCAGCAGAGATGACGAG 288
.....
83 ysrnAlaHisIleAspIleHisLys 91
```

seq\_name: /cgn2\_6/podata/2/1aa/5B\_COMB.pep:US-08-918-727-5

seq\_documentation\_block:

Sequence 5, Application US/08918727

Patent No. 5849528

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.

APPLICANT: Lal, Preeti

APPLICANT: Shah, Puri

TITLE OF INVENTION: HUMAN S100 PROTEINS

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/918,727

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0373 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 92 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Genbank

CLONE: 337730

US-08-918-727-5

alignment\_scores:

```
Quality: 177.50      Length: 88
Ratio: 2.536         Gaps: 1
Percent Similarity: 79.545      Percent Identity: 39.773
```

alignment\_block:

US-09-806-382a-2 x US-08-918-727-5 ..

Align seg 1/1 to: US-08-918-727-5 from: 1 to: 92

```
13 ATGTGCGAGCTGGAACGACATAGAGACCATCATCAACACTTCCACGA 62
|||||
1 MetSerGluLeuGluLysAlaMetValAlaLeuIleAspValPheHisgl 17
63 ATACTCTGAAGCTGGGGGCAACCCAGACCCCTGAACCGAGGGAATTCA 112
|||||
17 nTyrSerGlyArgGluGlyAspLysHisLysLeuLysLysSerGluLeu 34
```





```

113 AACAGCTGGTGCAGAAAGATCTGCAGAAATTTCTCAAGAGAGAAATAG 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 yseluleuileasnasngluLeuSerHisPheLeu...GIUGIUIleLys 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
163 AATGAAAGGTCTATAGACACATCATGAGGAGACCTGACCAATATGCAGA 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
50 GIUGIUGIuValValasPluysValMetGluThrLeuAspGluAspGlyAs 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
213 CAAGCAGTGCAGCTTCGAGGAGTTTCATCATGTCGATGCGAGGCTAACCT 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 pGIuGIuCIuCsAspPheGlnGluThrPheMetAlaPheValSerMetValThr 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
263 GGGCCTCCACGAG 276
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
83 hrAlaCySHISglu 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-987-272A-11

```

seq_documentation_block:
; Sequence 11, Application US/07987272A
; Patent No. 5731166
; GENERAL INFORMATION:
; APPLICANT: Ceczy, C., Simpson, R. J. and Lackmann, M
; TITLE OF INVENTION: No. 5731166el Chemolactic Factor
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman Darby & Cushman
; STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/987,272A
; FILING DATE: 05-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 2127
; FILING DATE: 05-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 4463
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Blinkman, David W
; REGISTRATION NUMBER: 20,817
; REFERENCE/DOCKET NUMBER: DWB/1925/200259
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861 3000
; TELEFAX: 202-822 0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-987-272A-11

```

```

alignment_scores:
Quality: 172.50      Length: 87
Ratio: 2.500         Gaps: 1
Percent Similarity: 79.310  Percent Identity: 39.080

```

alignment\_block:  
US-09-806-382A-2 x US-07-987-272A-11 ..

```

Align seq 1/1 to: US-07-987-272A-11 from: 1 to: 91
113 TCCGAGCTGAGAGCAACATAGAGACCATCATCAACCTTCCACCATATA 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 SerGIuLeuGIuLysAlaMetValAlaLeuIleAspAlaPheHisGlnTy 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 CTCTGTGAAAGCTGGGCGACCCAGACACACCTGAACACGAGGGGAATTCAAG 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 rSerGIuYArgGluGlyAspLysHisLysLysLysLysSerGIuLeuYsG 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
116 AGCTGTGCGAAAGATCTGCAGAAATTTCTCAAGAGAGAAATAGCAAT 165
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 IuLeuIleAsnAsnGluLeuSerHisPheLeu...GIUGIUIleLysGlu 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
166 GAAAGGTCTATAGACACATCATGAGGAGACCTGACCAATATGCAGCAA 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
50 GIUGIuValValasPluysValMetGluThrLeuAspGlnAspGlyAspGI 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 yGIuCIuCsAspPheGlnGluThrPheMetAlaPheValAlaMetValThr 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
266 CCTCCACGAG 276
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
83 lAcYSHISglu 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-662-198B-2

```

seq_documentation_block:
; Sequence 2, Application US/07662198B
; Patent No. 5262528
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Lee, Sam W.
; APPLICANT: Tomasetto, Catherine
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,198B
; FILING DATE: 19910228
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/049001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear
; US-07-662-198B-2

```

```

1 seq_documentation_block:
2 ; Sequence 2, Application US/08190560
3 ; Patent No. 5798257
4
5 GENERAL INFORMATION:
6
7 APPLICANT: Zain, Saeeda
8 APPLICANT: Luknaldin, Eugene
9 TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
10 TITLE OF INVENTION: THE MTS-1 GENE
11 NUMBER OF SEQUENCES: 9
12
13 CORRESPONDENCE ADDRESSES:
14 ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
15 STREET: 400 garden City Plaza
16 CITY: Garden City
17 STATE: New York
18 COUNTRY: United States
19
20 ZIP: 11530
21
22 COMPUTER READABLE FORM:
23 MEDIUM TYPE: Floppy disk
24 COMPUTER: IBM PC compatible
25 OPERATING SYSTEM: PC-DOS/MS-DOS
26 SOFTWARE: Patentn Release #1.0, Version #1.25
27
28 CURRENT APPLICATION DATA:
29 APPLICATION NUMBER: US/08/190,560
30 FILING DATE:
31
32 CLASSIFICATION: 435
33 ATTORNEY/AGENT INFORMATION:
34 NAME: Digiglio, Frank S.
35 REGISTRATION NUMBER: 31,346
36 REFERENCE/DOCKET NUMBER: 78792X
37 TELECOMMUNICATION INFORMATION:
38 TELEPHONE: (516) 742-4343
39 TELEFAX: (516) 742-4366
40
41 TELEX: 230 901 SANS UR
42
43 INFORMATION FOR SEO ID NO.: 2:
44
45 SEQUENCE CHARACTERISTICS:
46 LENGTH: 101 amino acids
47 TYPE: amino acid
48 STRANDEDNESS: single
49 TOPOLOGY: linear
50
51 MOLECULE TYPE: protein
52
53 US-08-190-560-2

```

```

1 ATGACTTCGAAATATGTCGACCTGGAGCAACATDAGACCATCATCA 50
  |||||
1 MetAlaCys.....ProLeuGluLysAlaLeuaspValMetValSe 14
51 CACCTTCACCAATATCTGTGTGAAGCTGGGGACCCAGACACCTGAAC 100
  :|||||:|||||  ||  ||  |||||
14 rThrPheHisLysTyrSerGlyLysGluGlyAspLysPheLysLeuSnl 31
101 AGGGGGAATTCAAAGAGCTGCTGGCAAAAGATCTGCAGAAATTTTCTAAG 150
  :|||:|||||:|||||  |||  |||
31 ySerGluLeuLysGluLeuLeuThrArgGluPheProSerPheLeuGly 47
151 AAGGAGATTAAGAAATGAAGGTGCTATACAAACAATTCATGGAGCACTGGA 200
  |||||
48 LysArgThr...AspGluAlaAlaPheGluLysLeuMetSerAsnLeuAs 63
201 CACAATATCGACAGCAAGCAGACTGAGCTTCGAGAGATTCATCTGCTGATGG 250
  ||:||||  |||||:|||||:|||||
63 pSerAsnArgAspAsnGluValAspPheGluGluTyrCysValPheLeuS 80
251 CGAGGCTTAACCTGGGCTTCACGACGAAGAATGCAGAGGTT.....GAC 294
  :|||:|||||  :|||  |||  |||
80 eCysAlaAlaLeuMetCysAsnGlu...PhePheGluGlyPheProAsp 95
295 GAGGGGCCCT 303
  :|||
96 LysGlnPro 98

```

100

100

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2002, 06:44:07 ; Search time 66.65 Seconds  
(without alignments)  
1271.471 Million cell updates/sec

Title: US-09-806-382A-2

Perfect score: 345  
Sequence: 1 atgaacttgcaaatgtcgcga.....tcggaggaggaacccctaa 345

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents.NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCBUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	345	100.0	571	1 US-08-385-241-4	Sequence 4, Appl1
2	197	57.1	4440	1 US-08-200-016-4	Sequence 4, Appl1
3	98.8	28.6	273	3 US-08-794-000-3	Sequence 3, Appl1
4	93.2	27.0	429	2 US-08-568-310D-1	Sequence 1, Appl1
5	93.2	27.0	429	4 US-09-270-455-1	Sequence 1, Appl1
6	77.4	22.4	441	2 US-08-568-310D-12	Sequence 12, Appl1
7	77.4	22.4	441	4 US-09-270-455-12	Sequence 12, Appl1
8	56.2	16.3	452	1 US-07-662-198B-1	Sequence 1, Appl1
9	56.2	16.3	452	1 US-08-322-742-1	Sequence 1, Appl1
10	47	13.6	579	1 US-08-190-560-3	Sequence 3, Appl1
11	47	13.6	579	1 US-08-469-277-3	Sequence 3, Appl1
12	47	13.6	579	2 US-08-468-946-3	Sequence 3, Appl1
13	47	13.6	579	2 US-08-468-942-3	Sequence 3, Appl1
14	46.6	13.5	10952	1 US-08-602-036A-1	Sequence 1, Appl1
15	46.6	13.5	10952	2 US-08-502-374A-1	Sequence 1, Appl1
16	46.6	13.5	10952	2 US-08-642-407A-1	Sequence 1, Appl1
17	46.4	13.4	1141	2 US-08-918-727-4	Sequence 4, Appl1
18	46.4	13.4	1141	3 US-09-205-680A-4	Sequence 4, Appl1
19	45	13.0	303	1 US-08-190-560-1	Sequence 1, Appl1
20	45	13.0	303	2 US-08-469-277-1	Sequence 1, Appl1
21	45	13.0	303	1 US-08-468-946-1	Sequence 1, Appl1
22	45	13.0	303	2 US-08-468-942-1	Sequence 1, Appl1
23	43.4	12.6	5552	3 US-08-155-888-1	Sequence 1, Appl1
24	43	12.5	347	4 US-09-328-111-848	Sequence 848, App
25	41.2	11.9	7218	1 US-08-232-463-14	Sequence 14, Appl1
26	41	11.9	408	1 US-08-385-241-2	Sequence 2, Appl1
27	41	11.9	418	1 US-07-987-272A-15	Sequence 15, Appl1

28	40	11.6	3176	1 US-08-425-299A-3	Sequence 3, Appl1
29	40	11.6	3176	2 US-08-486-663A-14	Sequence 14, Appl1
30	40	11.6	3176	2 US-08-247-904B-13	Sequence 13, Appl1
31	40	11.6	3176	3 US-08-767-942A-14	Sequence 14, Appl1
32	39.8	11.5	1080	2 US-08-918-727-2	Sequence 2, Appl1
33	39.8	11.5	1080	3 US-09-205-680A-2	Sequence 2, Appl1
34	39.6	11.5	433	1 US-07-987-272A-13	Sequence 13, Appl1
35	38.4	11.1	3176	2 US-09-070-060-10	Sequence 10, Appl1
36	38.4	11.1	3177	3 US-09-357-746-10	Sequence 10, Appl1
37	37.8	11.0	289	4 US-09-007-005-17	Sequence 17, Appl1
38	37.8	11.0	289	4 US-09-244-796-17	Sequence 17, Appl1
39	35	10.1	4195	1 US-08-200-016-1	Sequence 1, Appl1
40	34.2	9.9	1302	1 US-08-913-050A-2	Sequence 2, Appl1
41	34.2	9.9	1466	2 US-08-749-902-2	Sequence 2, Appl1
42	33.8	9.8	1114	3 US-09-048-889-10	Sequence 10, Appl1
43	33.6	9.7	2589	2 US-08-482-728A-3	Sequence 3, Appl1
44	33	9.6	1599	3 US-08-853-733B-1	Sequence 1, Appl1
45	33	9.6	1603	1 US-08-625-209A-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-08-385-241-4  
Sequence 4, Application US/08385241  
Patent No. 5776348  
GENERAL INFORMATION:  
APPLICANT: Selengut Ph.D., Jeremy D.  
APPLICANT: Orme-Johnson Ph.D., William H.  
APPLICANT: Dretler M.D., Stephen P.  
APPLICANT: Asakura M.D., Hirokaka  
TITLE OF INVENTION: SYSTEM AND METHOD FOR INHIBITING  
FORMATION OF CRYSTALLINE STRUCTURES THAT INCLUDE STRUTITE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Choate Hall & Stewart  
STREET: 53 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2891  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/385,241  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Herschbach Ph.D., Brenda M.  
REGISTRATION NUMBER: P-39,223  
REFERENCE/DOCKET NUMBER: 492611-000 (MIT6915)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-5175  
TELEFAX: (617) 248-4000  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 571 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
IMMEDIATE SOURCE:  
CLONE: hMRP-14 CDNA  
US-08-385-241-4  
Query Match 100.0%, Score 345, DB 1, Length 571;

Best Local Similarity 100.0%; Pred. No. 7.6e-92;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 atgacttcgaaatgctgcagctggaagcaacatagagaccatcatacaaccccttcac 60
DB 44 ATGACTTCGAAATGCTGCAGCTGGAGCGCAACATAGAGACCATTATCAACCTTCCAC 103
QY 61 caatcctctgaaagctgggagccagacaccctggaaccaggggaattcaaaagctg 120
DB 104 CAATCTCTGTGAAGCTGGGCGACCCAGACACCTTGAACCGAGGGGAATTCAAAAGCTG 163
QY 121 gtgcgaaagatctgcgaaatcttctcaagaaggaataaagatgaaagctcagaa 180
DB 164 GTGCAAAAGATCTGCAAAATTTTCTCAAGAAAGAGAAATGAAGAAAGGTATAGAA 223
QY 181 caccatcagggagagctgagacacaaatgacagcaagcgtgagcttgagagatc 240
DB 224 CACATCATGAGGAGCCTGAGCACAATGACAGCAAGCAGCTGAGCTTGAGAGATTCAATC 283
QY 241 atgctgtagcgaggttaacctggtgctcccgacgagaagatcacgaggtgtagcagagc 300
DB 284 ATGCTGATGGCGAGGCTTAACCTGGGCTCTCCACGAGAAATGACAGAGGTTACAGAGGC 343
QY 301 cctggcacaaccataagccaggcctcgaggaggacaccctaa 345
DB 344 CTGGCCACACACATAGCCAGGCTCGGGAGGCGACCCCTAA 388
```

## RESULT 2

US-08-200-016-4

Sequence 4, Application US/08200016

Patent No. 5614397

GENERAL INFORMATION:

APPLICANT: Weissman, Irving

APPLICANT: Lagasse, Eric

TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR MODULATING

TITLE OF INVENTION: APOPTOSIS IN HEMATOLOGY CELLS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish &amp; Richardson

STREET: 2200 Sand Hill Road, Suite 100

CITY: Menlo Park

STATE: California

COUNTRY: US

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/200.016

FILING DATE: 22-FEB-1994

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Rae-Venter, Barbara

REGISTRATION NUMBER: 32,750

REFERENCE/DOCKET NUMBER: 06037/003001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 854-5277

TELEFAX: (415) 854-0875

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 4440 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 2738

OTHER INFORMATION: /note= "Nucleotide indicated as "N"

OTHER INFORMATION: represents a region of 360 base pairs which do not

OTHER INFORMATION: appear in Figure 2."

US-08-200-016-4

Query Match 57.1%; Score 197; DB 1; Length 4440;  
Best Local Similarity 100.0%; Pred. No. 3.8e-48;  
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 149 aagaagagataagaatgaagaagtcataagaacacatcatgtagagagcctggacacaatg 208
DB 3447 AGAAGAGAGATGAAGATGAAGAGGTCATAGAACATCATGTGAGAGACCTGGACACAAATG 3506
QY 209 cagacaagcagctgagcttcgagaggttcatactgtagtggcagagctgagcct 268
DB 3507 CAGACAGAGCTGAGCTTCGAGAGGTTTCATCTGATGCGAGGCTTAACCTGGGCT 3566
QY 269 cccacgagaagatgcaagaggtgacagagccctggccacacataagcagcctg 328
DB 3567 CCCACGAGAGATGACAGCGGGGTGACGAGGCGCTGGCCACCATAGCCAGGCTTG 3626
QY 329 gggaggcgacccctaa 345
DB 3627 GGGAGGCGACCCCTAA 3643
```

## RESULT 3

US-08-794-000-3

Sequence 3, Application US/08794000

Patent No. 6087123

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Metal-Containing Ribonucleotide Polypeptides

NUMBER OF SEQUENCES: 4

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/794.000

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/DE96/01337

FILING DATE: 17-JUL-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 195 25 992.0

FILING DATE: 17-JUL-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 195 30 500.0

FILING DATE: 18-AUG-1995

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 273 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-794-000-3

Query Match 28.6%; Score 98.8; DB 3; Length 273;  
Best Local Similarity 63.0%; Pred. No. 5.3e-20;  
Matches 170; Conservative 0; Mismatches 97; Indels 3; Gaps 1;

```
QY 20 agctggaacgaactaagagacatcaaaccttcacccaataactctggaagctgg 79
DB 5 AGCTGGAGGACCACTGAGGGCGCATCATCATCTTCCACAGTACTGTGCGGCTGG 64
QY 80 ggcaccagacacccctggaacaggggaattcaaaccttggtgcaaaagatctgcaa 139
DB 65 GCCACTATGACACCTGTATCAAGCGGAGCTGATCAACAGAGAGCTGCCCA 124
QY 140 atttctcaagaagagagaataagaatgaagaagtcataagaacacatcatgagagcctgg 199
```

Db 125 ACACCT---GAAGACCCAAAGACAGGACACCATTTGACAAATCTTCCAGAACTCG 181  
Qy 200 aacaaatgcaagaagcagctgagctcgaagaatcattcgtatgagcgaagctaa 259  
Db 182 ATGCCAACGAGGATGAGCAGGATGACCTTCAAGAGATTGTGTGCTGTGATGACATGTGC 241  
Qy 260 cctgggctcccaagagaagatgacgaag 289  
Db 242 TGATCAGAGCCCATGACAAACATCCACAAG 271

RESULT 4  
US-08-568-310D-1  
; Sequence 1, Application US/08568310D  
; Patent No. 5976832  
; GENERAL INFORMATION:  
; APPLICANT: HITOMI, JIRO  
; APPLICANT: YAMAGUCHI, KEN  
; APPLICANT: YAMAMURA, TOKUJIRO  
; APPLICANT: KIMURA, TATSUJI  
; TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE  
; STREET: 99 PARK AVENUE  
; STREET: 6th FLOOR  
; CITY: NEW YORK CITY  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB  
; MEDIUM TYPE: STORAGE  
; COMPUTER: IBM-PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS 6.2  
; SOFTWARE: WORDPERFECT 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/568, 310D  
; FILING DATE: DECEMBER 6, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 7-70468 and 7-45564(both Japan)  
; FILING DATE: 3/6/95 and 3/6/95, respectively  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KLEIN, MILTON  
; REGISTRATION NUMBER: 27101  
; REFERENCE/DOCKET NUMBER: 3316  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)953-3350  
; TELEFAX: (212)953-3352  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 429  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; PUBLICATION INFORMATION:  
; RELEVANT RESIDUES IN SEQ ID NO: 1:  
; RELEVANT RESIDUES IN SEQ ID NO: FROM 1 TO 429  
US-08-568-310D-1

Query Match 27.0%; Score 93.2; DB 2; Length 429;  
Best Local Similarity 58.8%; Pred. No. 2, 9e-18;  
Matches 180; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

Qy 6 ttgcaaatgctgcagctggaagcaatagacacatcacaacctccaccaata 65  
Db 41 TGGTAAATGACTAAGCTGAGATCACCTGAGGGAATCATCAACATCTCCACCACTA 100  
Qy 66 cctctggaagcctggggcaccacacacctgacacagggggaattcaagagctgtg 125

Db 101 CTCGGTTCCGGTGGGGCATTTTCACACCTCTCAACAGCGGTGAGCTGACCTGATTCAC 160  
Qy 126 aaaagatctgcaaaatttctcaagaagagaaatgaatgaagaaagtcataagacacat 185  
Db 161 AAGGAACCTTCCCAAAACCTCCAGA---ACACCAAAAGATCAACCTACCATTCACAAAT 217  
Qy 186 catggagagacctggacaacaatgcaagaagcagctgagcttcgaggaattcatgct 245  
Db 218 ATTCCAGAGCTGGAATGCCGATAAAGACGAGCGGTGACCTTTGAGGAATTCGTACTCT 277  
Qy 246 gatggcagagctaacctgggctccacgagaagatgcaagaggtgacagagggccctgg 305  
Db 278 GGTGTCAGGGGTCTGTAAGAACAGCCACATAGATATCCAAAGAGTAGGAAGCTCTTTC 337  
Qy 306 ccacca 311  
Db 338 CAGCAA 343

RESULT 5  
US-09-270-455-1  
; Sequence 1, Application US/09270455  
; Patent No. 6313267  
; GENERAL INFORMATION:  
; APPLICANT: HITOMI, JIRO  
; APPLICANT: YAMAGUCHI, KEN  
; APPLICANT: YAMAMURA, TOKUJIRO  
; APPLICANT: KIMURA, TATSUJI  
; TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE  
; STREET: 99 PARK AVENUE  
; STREET: 6th FLOOR  
; CITY: NEW YORK CITY  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB  
; MEDIUM TYPE: STORAGE  
; COMPUTER: IBM-PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS 6.2  
; SOFTWARE: WORDPERFECT 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/270, 455  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/568, 310  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KLEIN, MILTON  
; REGISTRATION NUMBER: 27101  
; REFERENCE/DOCKET NUMBER: 3316  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)953-3350  
; TELEFAX: (212)953-3352  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 429  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; PUBLICATION INFORMATION:  
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 429  
US-09-270-455-1

Query Match 27.0%; Score 93.2; DB 4; Length 429;  
Best Local Similarity 58.8%; Pred. No. 2, 9e-18;





```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 12: FROM 1 TO 441
US-09-270-455-12

```

```

Query Match      22.4%; Score 77.4; DB 4; Length 441;
Best Local Similarity 56.8%; Pred. No. 1.2e-13;
Matches 163; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

```

```

QY 3 gacttgcacaaagtgcacgtcagcagcaacatagagaccatcatcaaacctccacca 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12 GCGTGGAGATGACAAACCTTGAAAGACATCTGAGGAAATTTGCAATATCTCCACCA 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 63 atactctgtaagctggtgagcaccacacacccctgaaacaggggaattcaaaagctgt 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 72 ATACTCAGTTGCGAAGGGGCAATTTTGACACCCCTCTTAAGGGTGAAGTGAACAGCTGCT 131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 ggcgaagaatctgcaaatcttctcaagaagagaaatgaatgaatgaatgaatga 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 132 TACAAGGACCTTGCAAAACACATCAAGA--ATATCAAGAATAAAGCTGTGATGATGA 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 183 calcatggaagacctggaacacaatgcagaacagcagctgagcttcgagagatccat 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 AATATTCGAAGGCTGTGATCTATCAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGA 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 243 gctgatggcgaagctaacctcgtgacctcccaagagaatgcagcagc 289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 249 CCTGTGATGCCATTGCGCTGAAGGCTGCCATTCACACACCACCAAG 295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

# RESULT 8

```

US-07-662-198B-1
; Sequence 1, Application US/07662198B
; Patent No. 5262528
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Lee, Sam W.
; APPLICANT: Tomasello, Catherine
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/662,198B
FILING DATE: 19910228
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/049001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154

```

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 452
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
US-07-662-198B-1

```

```

Query Match      16.3%; Score 56.2; DB 1; Length 452;
Best Local Similarity 58.0%; Pred. No. 1.9e-07;
Matches 119; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

```

```

QY 47 tcaaaccttcaccaactctgtgaagctggtgagccccaagacaccttgaacagggg 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 85 TCACATACCTTCCAAAGATCTCTGCAAGAGGCGCAAGTTCAAGCTTGATGAAGGGGG 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 107 aattcaaaagctgtgtggaagaatctgcaaaatttctcaagaagaatgaatga 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 145 AATTAAGGAATCTCTGCAAGAGAGCTGCCACCTTTGT---GGGGGAGAAAGTGAATG 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 167 aaaaagtcatagaacacatcatggaagccttgacacacaatgcagaagaagctgagct 226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 202 AGAAGGGGCTGAAAGAGCTGATGGCAACCTGATGAGACAGTGAACCAAGTGAAGTGA 261
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 227 tcgagagcttcacatgctgagtgc 251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 262 TCAGAGAGATGCTGTTTCTCTGCG 286
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

# RESULT 9

```

US-08-322-742-1
; Sequence 1, Application US/08322742
; Patent No. 5688641
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,742
FILING DATE:

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/938,823
FILING DATE: September 1, 1992
APPLICATION NUMBER: 07/844,296
FILING DATE: February 28, 1992
APPLICATION NUMBER: 07/552,216
FILING DATE: February 28, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/048003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154

```

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 452
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-322-742-1.

```

Query Match 16.3%; Score 56.2; DB 1; Length 452;  
Best Local Similarity 58.0%; Pred. No. 1.9e-07;  
Matches 119; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

QY 47 tcaaaccttcaccaatactctgtgaagctgggacaccagacacctgaaccaggag 106  
DB 85 TCACATCCTTCACAAAGTACTCTCCCAAGAGGCGACAACTCAAGCTGAGTAAGGGG 144  
QY 107 aattcaagaagctgtgggaaaaagctgcgaattttctcaagaaggaataagatg 166  
DB 145 AATGAGGAACTTCTGACAAAGAGCTGCCAGCTTGT--GGGGAGAAAGTGGATG 201  
QY 167 aaaaagtcataagacacatcatgagagacctgtgacacaaatgcagaagaagctgagct 226  
DB 202 AGAGAGGAGCTGAAGAAGCTGAGGCAACCTGATGATGAACAGTGAACAGAGTGGAGCT 261  
QY 227 tcgaggagctcatcaltgctgtaggc 251  
DB 262 TCCAGAGATGCTGTCTTTCTCGGC 286

## RESULT 10

US-08-190-560-3  
; Sequence 3, Application US/08190560  
; Patent No. 5798257  
; GENERAL INFORMATION:  
; APPLICANT: Zain, Sayeeda  
; APPLICANT: Lukanidin, Eugene  
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY  
; TITLE OF INVENTION: THE MTS-1 GENE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190,560  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DIGILIO, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 78792X  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 579 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
US-08-190-560-3

Query Match 13.6%; Score 47; DB 1; Length 579;  
Best Local Similarity 50.9%; Pred. No. 0.00011;  
Matches 138; Conservative 0; Mismatches 130; Indels 3; Gaps 1;  
QY 5 ctgcaaatgtcgcagctgtgaagcacaatagagacatcatcaaaccttcacacaa 64  
DB 131 CTGTATGCGCGTCCCTGTGAGAAAGGCCCTGATGTGATGTCTCCACCCTTCACAAAGT 190

DB 131 CTGTATGCGCGTCCCTGTGAGAAAGGCCCTGATGTGATGTCTCCACCCTTCACAAAGT 190  
QY 65 actctgtgaagctggggagaccaccagacccctgaaccaggagggaattcaagagctgtgc 124  
DB 191 ACTCGGCAAGAGGGGTACAGTTCAAGCTCAACAAAGTCAAGAACTAAAGAGCTCTGA 250  
QY 125 gaaaagatctgaaattttctcaagaaggaataaagaatgaagaagtcataagaacaa 184  
DB 251 CCCGGAGACTGCGCCAGCTTCTTGGGAAAGACAG--ATGAAGCTGCTTCCAGAGC 307  
QY 185 tcatgagagacttgagacaaatgcagacagcagctgcgagcttcgagagatcatcagc 244  
DB 308 TATATGCAACTTGGACGCAACAGGCAACGAGGTGAGACTTCCAAAGAGTACTGTCT 367  
QY 245 tgatgcagagctaacctctggcctccacga 275  
DB 368 TCTGTCTGATGATGCGCATGATGTGTAAAGCA 398

## RESULT 11

US-08-469-277-3  
; Sequence 3, Application US/08469277  
; Patent No. 5801142  
; GENERAL INFORMATION:  
; APPLICANT: Zain, Sayeeda  
; APPLICANT: Lukanidin, Eugene  
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY  
; TITLE OF INVENTION: THE MTS-1 GENE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,277  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190,560  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DIGILIO, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 78792X  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 579 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
US-08-469-277-3

Query Match 13.6%; Score 47; DB 1; Length 579;  
Best Local Similarity 50.9%; Pred. No. 0.00011;  
Matches 138; Conservative 0; Mismatches 130; Indels 3; Gaps 1;  
QY 5 ctgcaaatgtcgcagctgtgaagcacaatagagacatcatcaaaccttcacacaa 64  
DB 131 CTGTATGCGCGTCCCTGTGAGAAAGGCCCTGATGTGATGTCTCCACCCTTCACAAAGT 190





Tue Sep, 10 09:48:16 2002

us-09-806-382a-2.rni

Page 9

Db 2387 CACCTGGACCCCGGTAGCGACGTACAGATGTCCTCCCTCTCTCTCCGGGGGAGGCACCT 2446  
 QY 198 g 198  
 Db 2447 G 2447

Search completed: September 10, 2002, 06:44:15  
Job time: 4961 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: September 10, 2002, 06:49:27 ; Search time 298.43 Seconds  
(without alignments)  
1984.839 Million cell updates/sec

Title: US-09-806-382A-2

Perfect score: 345  
Sequence: 1 atgacttgcaaaatgctgca.....tcggggagggcaccctctaa 345

## Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

N\_Geneseq\_032802.\*  
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	345	100.0	345	21 AAC81809	Human S100A9 cDNA.
2	345	100.0	345	21 AA11970	Human calcium-bind
3	345	100.0	565	22 AAF54714	Nucleotide sequenc
4	345	100.0	565	22 AAF54719	Nucleotide sequenc
5	345	100.0	571	19 AAV34698	Human caliproteclin
6	345	100.0	687	20 AA233673	Human breast tumou
7	345	100.0	688	20 AA233615	Human breast tumou
8	345	100.0	771	22 AA159150	Human polynucleoti
9	267.4	77.5	452	22 AAF54717	Nucleotide sequenc

10	262.2	76.0	342	22 AAF54709	Nucleotide sequenc
11	261.6	75.8	345	21 AA160936	Human polynucleoti
12	260.2	75.4	385	21 AAC06727	Human secreted pro
13	242	70.1	430	22 AAF54715	Nucleotide sequenc
14	241	69.9	425	22 AAF54713	Nucleotide sequenc
15	197	57.1	592	22 ABA59810	Human foetal liver
16	197	57.1	592	22 ABA28288	Probe #6754 for ge
17	197	57.1	592	22 AAK08083	Human brain expres
18	197	57.1	592	22 AAK33958	Human bone marrow
19	197	57.1	592	22 AA116584	Probe #6517 for ge
20	197	57.1	592	22 AA139682	Probe #8368 used t
21	197	57.1	4439	22 AAF54718	Nucleotide sequenc
22	197	57.1	4440	18 AAT68322	Human multidrug re
23	195	56.5	195	22 ABA72351	Human foetal liver
24	195	56.5	195	22 ABA38167	Probe #16633 for g
25	195	56.5	195	22 AAK20777	Human brain expres
26	195	56.5	195	22 AAK46926	Human bone marrow
27	195	56.5	195	22 AA125780	Probe #15713 for g
28	195	56.5	195	22 AA152763	Probe #21449 used
29	181.6	52.6	406	22 AAF54712	Nucleotide sequenc
30	177.4	51.4	350	20 AAX23464	Human neutrophil c
31	150	43.5	594	22 ABA61147	Human foetal liver
32	150	43.5	594	22 ABA29027	Probe #7493 for ge
33	150	43.5	594	22 AAK09442	Human brain expres
34	150	43.5	594	22 AAK35334	Human bone marrow
35	150	43.5	594	22 AA116996	Probe #6829 for ge
36	150	43.5	594	22 AA141047	Probe #9733 used t
37	133	38.6	133	22 ABA73673	Human foetal liver
38	133	38.6	133	22 ABA38897	Probe #17363 for g
39	133	38.6	133	22 AAK22120	Human brain expres
40	133	38.6	133	22 AAK48287	Human bone marrow
41	133	38.6	133	22 AA126189	Probe #16122 for g
42	133	38.6	133	22 AA154112	Probe #22798 used
43	109.6	31.8	477	22 AAF54711	Nucleotide sequenc
44	109	31.6	305	22 AAF54716	Nucleotide sequenc
45	103.2	29.9	347	23 AAS96258	Human bone marrow

## ALIGNMENTS

RESULT 1	
ID AAC81809	standard; cDNA; 345 BP.
XX	
XX AAC81809;	
XX	
DT 22-FEB-2001	(first entry)
XX	
DE Human S100A9 cDNA.	
XX	
KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;	
KW calcium-binding protein; calcium homeostasis; cardiac muscle;	
KW pumping capacity; myocardial cell; systolic calcium ion release;	
KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;	
KW valve defect; ss.	
XX	
OS Homo sapiens.	
XX	
PN DE19915485-A1.	
PD 19-OCT-2000.	
XX	
PF 07-APR-1999;	99DE-1015485.
XX	
PR 07-APR-1999;	99DE-1015485.
XX	
PA (KATU/) KATUS H A.	
XX (REMP/) REMPPIS A.	
XX	
PI Katus HA, Remppis A;	
XX	
DR WPI; 2000-673510/66.	

DR P-PSDB: AAB45539.  
XX Composition containing S100 protein, corresponding nucleic acid or  
PT vector, useful for treating cardiomyopathy and cardiac insufficiency -  
XX  
PS Claim 36; Page 16; 36pp; German.  
XX  
CC This invention describes a novel composition for treating primary or  
CC secondary cardiomyopathy or cardiac insufficiency contains at least one  
CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or  
CC fragments, or a gene transfer vector containing (II), optionally  
CC formulated with auxilliaries and/or carriers. (I) are calcium-binding  
CC proteins involved in calcium homeostasis, so their overexpression in  
CC cardiac muscle will improve pumping capacity (and overall contraction) of  
CC the heart. In cultured myocardial cells they increase the contraction and  
CC relaxation rates associated with increased systolic calcium ion release  
CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are  
CC used to treat cardiomyopathy (CMP) where inherited or caused by  
CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,  
CC dilative CMP caused by toxic/infectious disease, cardiac disease caused  
CC by pulmonary and/or arterial hypertension, and structural disease caused  
CC by rhythm disorders or valve defects, generally any condition associated  
CC with reduced contractile force. Unlike calmodulin, which is expressed  
CC ubiquitously, (I) show tissue-specific expression and treat the  
CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac  
CC disease.  
XX  
SQ Sequence 345 BP; 110 A; 89 C; 92 G; 54 T; 0 other:  
XX  
Query Match 100.0%; Score 345; DB 21; Length 345;  
Best Local Similarity 100.0%; Pred. No. 2e-86;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgacttgcaaaatgtcgcagctggaagcgaacatagagaccatcataacaccccttcac 60  
DB 1 atgacttgcaaaatgtcgcagctggaagcgaacatagagaccatcataacaccccttcac 60  
QY 61 caatactctgtgaagctgtggcaccacacccctgaaccagggggaattcaaaagctg 120  
DB 61 caatactctgtgaagctgtggcaccacacccctgaaccagggggaattcaaaagctg 120  
QY 121 gtgcgaaaagaatctgcgaatttctcaagaagagaataagaatgaagaaggtcatagaa 180  
DB 121 gtgcgaaaagaatctgcgaatttctcaagaagagaataagaatgaagaaggtcatagaa 180  
QY 181 cacatcatgagagacctgacacaaatgcagacagcgcgtgaggttcgagagttcatc 240  
DB 181 cacatcatgagagacctgacacaaatgcagacagcgcgtgaggttcgagagttcatc 240  
QY 241 atgctgatgagcgtactaaccttggtcctccacagagaagatgcacgaggtgacgaagggc 300  
DB 241 atgctgatgagcgtactaaccttggtcctccacagagaagatgcacgaggtgacgaagggc 300  
QY 301 cctggcaccacccataagccagcgtctcgggagggacccctctaa 345  
DB 301 cctggcaccacccataagccagcgtctcgggagggacccctctaa 345  
RESULT 2  
AA11970  
ID AA11970 standard; DNA: 345 BP.  
XX  
AC AA11970:  
XX  
XX 04-AUG-2000 (first entry)  
DE Human calcium-binding protein encoding DNA #2.  
XX  
XX Calcium-binding protein; granule release; calgranulin; human;  
KW vascular membrane growth; adult respiratory distress syndrome;  
KM acute myocardial infarction; ischemic reperfusion disorder;  
KM glomerulonephritis; rheumatoid arthritis; chronic bronchitis;

KW cerebral vascular disorder; asthma; peripheral circulation disturbance;  
KM angina pectoris; hypertension; multiple sclerosis; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200018970-A1.  
XX  
XX 06-APR-2000.  
XX  
XX 28-SEP-1999; 99WO-JP05302.  
XX  
XX 29-SEP-1998; 98JP-0274574.  
XX  
XX (ASAH ) ASAMI KASEI KOGYO KK.  
PI Seto M, Fukuda K;  
XX  
XX WPI; 2000-293189/25.  
DR P-PSDB: AAV87637.  
XX  
PT Controlling the release of granules from cell system using activated  
PT calgranulin for screening substances for granule activating or  
PT inhibiting activity -  
PS Claim 3(11); Page 39-40; 42pp; Japanese.  
XX  
CC This invention describes a novel method for controlling the release of  
CC granules by treating a cell system to increase or decrease activated  
CC calgranulin to enhance or depress the release of granules. Calgranulin  
CC is a calcium binding protein and can be used for controlling the release  
CC of granules from a cell system e.g. those involved with the inhibition of  
CC vascular membrane growth. Vascular membrane growth is associated with  
CC e.g. adult respiratory distress syndrome, acute myocardial infarction due  
CC to ischemic reperfusion disorders, glomerulonephritis, rheumatoid  
CC arthritis, chronic bronchitis, cerebral vascular disorders, asthma,  
CC peripheral circulation disturbance, angina pectoris, hypertension and  
CC multiple sclerosis. The new method is used for screening substances for  
CC their ability to activate or inhibit the release of granules. This  
CC sequence encodes a human calcium-binding protein which is described in  
CC the method of the invention.  
XX  
SQ Sequence 345 BP; 110 A; 89 C; 92 G; 54 T; 0 other:  
XX  
Query Match 100.0%; Score 345; DB 21; Length 345;  
Best Local Similarity 100.0%; Pred. No. 2e-86;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgacttgcaaaatgtcgcagctggaagcgaacatagagaccatcataacaccccttcac 60  
DB 1 atgacttgcaaaatgtcgcagctggaagcgaacatagagaccatcataacaccccttcac 60  
QY 61 caatactctgtgaagctgtggcaccacacccctgaaccagggggaattcaaaagctg 120  
DB 61 caatactctgtgaagctgtggcaccacacccctgaaccagggggaattcaaaagctg 120  
QY 121 gtgcgaaaagaatctgcgaatttctcaagaagagaataagaatgaagaaggtcatagaa 180  
DB 121 gtgcgaaaagaatctgcgaatttctcaagaagagaataagaatgaagaaggtcatagaa 180  
QY 181 cacatcatgagagacctgacacaaatgcagacagcgcgtgaggttcgagagttcatc 240  
DB 181 cacatcatgagagacctgacacaaatgcagacagcgcgtgaggttcgagagttcatc 240  
QY 241 atgctgatgagcgtactaaccttggtcctccacagagaagatgcacgaggtgacgaagggc 300  
DB 241 atgctgatgagcgtactaaccttggtcctccacagagaagatgcacgaggtgacgaagggc 300  
QY 301 cctggcaccacccataagccagcgtctcgggagggacccctctaa 345  
DB 301 cctggcaccacccataagccagcgtctcgggagggacccctctaa 345



Query Match	Best Local Similarity	Conservative	Mismatches	Indels	Gaps
1 atgatttcgaatgctgcagctgtgacgcaacacatagagaccatcatcaacaccttcac	100.0%	100.0%	0	0	0
32 atgatttcgaatgctgcagctgtgacgcaacacatagagaccatcatcaacaccttcac	100.0%	100.0%	0	0	0
61 caatactctgtgaaagcttgggacccacagacaccttgaacacgaggggaattcaaaagctg	100.0%	100.0%	0	0	0
92 caatactctgtgaaagcttgggacccacagacaccttgaacacgaggggaattcaaaagctg	100.0%	100.0%	0	0	0
121 gtgcgaaaaagatctgcataaatttctcaagaagaagaataagaatgaaaggtcatagaa	100.0%	100.0%	0	0	0
152 gtgcgaaaaagatctgcataaatttctcaagaagaagaataagaatgaaaggtcatagaa	100.0%	100.0%	0	0	0

Oy	181	caacatcgtgaagacctgtgacacacaatgcagacaagcagcttgaggttcatc	240
Db	212	cacatcttgagagacctgtgacacacaatgcagacaagcagcttgaggttcatc	271
Oy	241	atgctgatgtgcgaagcctaaccttggtcctccacgaanaatgcagaggttgacgaagcc	300
Db	272	atgctgatgtgcgaagcctaaccttggtcctccacgaanaatgcagaggttgacgaagcc	331
Oy	301	ctctgcccacacataaagccagagccttcggggaaggaacccctaa	345
Db	332	ctctgcccacacataaagccagagccttcggggaaggaacccctaa	376
RESULT 4			
AF54719	AAAF54719 standard; DNA; 565 BP.		
AC	AAAF54719;		
XX			
XX	15-MAY-2001 (first entry)		
DE			
XX	Nucleotide sequence of a human polynucleotide sequence.		
XX			
XX	Human: perlecan; retinol-binding plasma protein; calgranulin B; vaccine;		
KW	ganglioside GM2 activator; saposin B; degenerative disease; glial cell;		
KW	neurological disease; auto-immune disease; multiple sclerosis; toxicity;		
KW	Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;		
KW	rheumatoid polyarthritits; lupus erythematosus; gene therapy; ss.		
OS	Homo sapiens.		
XX			
PN	WO200105422-A2.		
XX			
PD	25-JAN-2001.		
XX			
PF	17-JUL-2000; 2000MO-FR02057.		
XX			
PR	15-JUL-1999; 99FR-0009372.		
XX			
PA	(INMR ) BIOMERIEUX STELHYS.		
XX			
PI	Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;		
DR	WPI: 2001-159475/16.		
XX			
PT	Detecting, preventing and treating degenerative, neurological and		
PT	autoimmune diseases, particularly multiple sclerosis, using specified		
PT	polypeptides or related nucleic acid or ligand		
XX			
PS	Claim 11; Page 188; 209pp; French.		
XX			
CC	The present sequence represents a human polynucleotide sequence, which		
CC	is used in the method of the invention. The specification describes a		
CC	method which uses at least one polypeptide or polynucleotide sequence		
CC	belonging to the perlecan, precursor of the retinol-binding plasma		
CC	protein, precursor of the ganglioside GM2 activator, calgranulin B or		
CC	saposin B protein families. The method is used for detecting,		
CC	preventing or treating a degenerative, neurological and/or auto-immune		
CC	disease. The polynucleotides and polypeptides are used for diagnosis,		
CC	prognosis, prevention and treatment of multiple sclerosis (in its		
CC	various forms and phases). They may also be useful in cases of		
CC	e.g. Alzheimer's and Parkinson's diseases, amyotrophic lateral sclerosis,		
CC	rheumatoid polyarthritits and lupus erythematosus, including use as		
CC	vaccines and in gene therapy (expression of sense or antisense		
CC	sequences). They can also be used to assess efficacy of potential		
CC	therapeutic agents, particularly compounds that reduce or inhibit		
CC	toxicity towards glial cells.		
XX			
SO	Sequence 565 BP; 158 A; 158 C; 156 G; 93 T; 0 other;		

Query Match
100.0%; Score 345; DB 22; Length 565;

```

Best Local Similarity 100.0%; Pred. No. 2.5e-86;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgacttgcaaaatgtcgagctggaacgcaacatagagacatctcaacacctccac 60
DB 32 atgacttgcaaaatgtcgagctggaacgcaacatagagacatctcaacacctccac 91
QY 61 caatactctgtlaagctgggacaccgacacctgaaccagggggaattcaaaagctg 120
DB 92 caatactctgtlaagctgggacaccgacacctgaaccagggggaattcaaaagctg 151
QY 121 gtgcgaaaaagatctgcaaaatttctcaagaaggagaataaagaatgacataga 180
DB 152 gtgcgaaaaagatctgcaaaatttctcaagaaggagaataaagaatgacataga 211
QY 181 caatacttgaggacctggaacaaatgagacaagcaactgagcttgaggagttcacc 240
DB 212 caatacttgaggacctggaacaaatgagacaagcaactgagcttgaggagttcacc 271
QY 241 atgctgatgtgagagctgaacctgggacctccacgagaagaatcacgaggtgacgagggc 300
DB 272 atgctgatgtgagagctgaacctgggacctccacgagaagaatcacgaggtgacgagggc 331
QY 301 cctggcaccaccacataagcgaagcctcggggaggggacccctaa 345
DB 332 cctggcaccaccacataagcgaagcctcggggaggggacccctaa 376

RESULT 5
AAV34698
ID AAV34698 standard; cDNA; 571 BP.
AC AAV34698;
XX
XX 03-SEP-1998 (first entry)
DT
XX Human calprotectin subunit MRP-14 protein encoding cDNA.
DE
XX Human; MRP-8; MRP-14; calprotectin; mineral precipitate; struvite;
KW calcium phosphate; kidney stone; renal calculi; struvite stone;
XX urinary tract infection; ds.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT 44..385
XX FT /*tag= a
XX FT /product= "human MRP-14"
XX
XX US5776348-A.
XX
XX 07-JUL-1998.
XX
XX 07-FEB-1995; 95US-0385241.
XX
XX 07-FEB-1995; 95US-0385241.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Asakura H, Dretler SP, Orme-johnson WH, Selengut JD;
XX WPI; 1998-397914/34.
XX
XX P-PSDB; AAW60178.
XX
XX Inhibiting kidney stone formation - uses the protein calprotectin
XX
XX Example 1; Columns 23-24; 19pp; English.
XX
XX This cDNA encodes a human calprotectin subunit MRP-14. This is used in
XX a method for inhibiting the formation of a mineral precipitate in a
XX solution which comprises providing a solution comprising component ions
XX of the mineral precipitate or its precursors and contacting the solution
XX with an effective amount of isolated calprotectin, or a derivative of it.

```

The method is useful for the inhibition of kidney stone formation (renal calculi). Kidney stones are concentrations of inorganic and organic salts that develop though crystal nucleation, aggregation and growth in the kidneys, which then can block the urether and if not passed to the bladder, grow and become symptomatic. Struvite stones (MgNH<sub>4</sub>PO<sub>4</sub>) are commonly found after urinary tract infection. The mechanisms of this and other stone formation is unclear, but some urine compositions are known to inhibit formation. One of the major components of these is a protein factor calprotectin, which inhibits mineral precipitation. Calprotectin can also be used to raise antibodies, which in turn can be used to detect the protein in samples. The levels of calprotectin found in samples can be compared to levels found in normal humans, and thus assuming increased calprotectin correlates to indication of kidney stone formation, this procedure can be used as a diagnostic tool.

Sequence 571 BP; 160 A; 160 C; 157 G; 94 T; 0 other;

```

Query Match          100.0%; Score 345; DB 19; Length 571;
Best Local Similarity 100.0%; Pred. No. 2.5e-86;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgacttgcaaaatgtcgagctggaacgcaacatagagacatctcaacacctccac 60
DB 44 atgacttgcaaaatgtcgagctggaacgcaacatagagacatctcaacacctccac 103
QY 61 caatactctgtlaagctgggacaccgacacctgaaccagggggaattcaaaagctg 120
DB 104 caatactctgtlaagctgggacaccgacacctgaaccagggggaattcaaaagctg 163
QY 121 gtgcgaaaaagatctgcaaaatttctcaagaaggagaataaagaatgacataga 180
DB 164 gtgcgaaaaagatctgcaaaatttctcaagaaggagaataaagaatgacataga 223
QY 181 caatacttgaggacctggaacaaatgagacaagcaactgagcttgaggagttcacc 240
DB 224 caatacttgaggacctggaacaaatgagacaagcaactgagcttgaggagttcacc 283
QY 241 atgctgatgtgagagctgaacctgggacctccacgagaagaatcacgaggtgacgagggc 300
DB 284 atgctgatgtgagagctgaacctgggacctccacgagaagaatcacgaggtgacgagggc 343
QY 301 cctggcaccaccacataagcgaagcctcggggaggggacccctaa 345
DB 344 cctggcaccaccacataagcgaagcctcggggaggggacccctaa 388

RESULT 6
AAZ33673
ID AAZ33673 standard; cDNA; 687 BP.
AC AAZ33673;
XX
XX 08-DEC-1999 (first entry)
DT
XX Human breast tumour-associated EST 63.
DE
XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;
XX treatment; tumour; cytostatic; medicament; ss.
XX
XX Homo sapiens.
XX
XX DE19813839-A1.
XX
XX 23-SEP-1999.
XX
XX 20-MAR-1998; 98DE-1013839.
XX
XX 20-MAR-1998; 98DE-1013839.
XX
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

```

XX WPI; 1999-528981/45.  
DR P-PSDB; AAY48615.  
XX  
XX Human nucleic acid sequences and protein products from tumor breast  
PT tissue, useful for breast cancer therapy -  
XX  
XX Claim 3; 141; 188bp; German.  
XX  
XX This invention describes novel human nucleic acid sequences from tumor  
CC breast tissue which have cytotatic activity. The nucleic acid sequences  
CC can be used to produce and isolate full-length gene sequences. They can  
CC be used to express proteins, which can be used as tools to find an  
CC activity against breast cancer. The sequences can be used in sense or  
CC antisense form. They are especially useful for medicaments for gene  
CC therapy to treat breast cancer. AA233611-248617 represents expressed  
CC sequence tags described in the method of the invention.  
XX  
XX Sequence 687 BP; 177 A; 198 C; 190 G; 122 T; 0 other;

Query Match 100.0%; Score 345; DB 20; Length 687;  
Best Local Similarity 100.0%; Pred. No. 2.7e-86;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgacttgcaaaatgtcgcagctggaacgcaacatagagaccatcatcaacaccttcac 60  
DB 101 atgacttgcaaaatgtcgcagctggaacgcaacatagagaccatcatcaacaccttcac 160  
QY 61 caatactctgtgaagctggggcaccagacacaccttgaaacagggggaattcaagaagctg 120  
DB 161 caatactctgtgaagctggggcaccagacacaccttgaaacagggggaattcaagaagctg 220  
QY 121 gtgcgaaaagatctgcgcaaatcttctcaagaagagagataagaatgaagaagctagaa 180  
DB 221 gtgcgaaaagatctgcgcaaatcttctcaagaagagagataagaatgaagaagctagaa 280  
QY 221 gtgcgaaaagatctgcgcaaatcttctcaagaagagagataagaatgaagaagctagaa 280  
DB 281 caatcatctggaagctggagacacaatgcaagcaagcagcttcgagagattcatc 340  
QY 241 atgctgtagtggaagctgaacctgggctccacagagaagatgacagagggtagcagagggc 300  
DB 341 atgctgtagtggaagctgaacctgggctccacagagaagatgacagagggtagcagagggc 400  
QY 301 cctggcaccaccataagccagcctcgggagggcaccacctaa 345  
DB 401 cctggcaccaccataagccagcctcgggagggcaccacctaa 445

RESULT 7  
AA233615  
ID AA233615 standard; cDNA; 688 BP.  
XX  
XX AA233615;  
XX  
XX 08-DEC-1999 (first entry)  
XX  
XX Human breast tumour-associated EST 5.  
XX  
XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;  
KW treatment; tumour; cytostatic; medicament; ss.  
XX  
XX Homo sapiens.  
XX  
XX DE19813839-A1.  
XX  
XX 23-SEP-1999.  
XX  
XX 20-MAR-1998; 98DE-1013839.  
XX  
XX 20-MAR-1998; 98DE-1013839.  
XX  
XX 20-MAR-1998; 98DE-1013839.  
XX

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
XX  
XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;  
PI  
XX WPI; 1999-528981/45.  
XX  
XX Human nucleic acid sequences and protein products from tumor breast  
PT tissue, useful for breast cancer therapy -  
XX  
XX Claim 3; 86; 188bp; German.  
XX  
XX This invention describes novel human nucleic acid sequences from tumor  
CC breast tissue which have cytotatic activity. The nucleic acid sequences  
CC can be used to produce and isolate full-length gene sequences. They can  
CC be used to express proteins, which can be used as tools to find an  
CC activity against breast cancer. The sequences can be used in sense or  
CC antisense form. They are especially useful for medicaments for gene  
CC therapy to treat breast cancer. AA233611-248617 represents expressed  
CC sequence tags described in the method of the invention.  
XX  
XX Sequence 688 BP; 178 A; 198 C; 190 G; 122 T; 0 other;

Query Match 100.0%; Score 345; DB 20; Length 688;  
Best Local Similarity 100.0%; Pred. No. 2.7e-86;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgacttgcaaaatgtcgcagctggaacgcaacatagagaccatcatcaacaccttcac 60  
DB 101 atgacttgcaaaatgtcgcagctggaacgcaacatagagaccatcatcaacaccttcac 160  
QY 61 caatactctgtgaagctggggcaccagacacaccttgaaacagggggaattcaagaagctg 120  
DB 161 caatactctgtgaagctggggcaccagacacaccttgaaacagggggaattcaagaagctg 220  
QY 121 gtgcgaaaagatctgcgcaaatcttctcaagaagagagataagaatgaagaagctagaa 180  
DB 221 gtgcgaaaagatctgcgcaaatcttctcaagaagagagataagaatgaagaagctagaa 280  
QY 221 gtgcgaaaagatctgcgcaaatcttctcaagaagagagataagaatgaagaagctagaa 280  
DB 281 caatcatctggaagctggagacacaatgcaagcaagcagcttcgagagattcatc 340  
QY 241 atgctgtagtggaagctgaacctgggctccacagagaagatgacagagggtagcagagggc 300  
DB 341 atgctgtagtggaagctgaacctgggctccacagagaagatgacagagggtagcagagggc 400  
QY 301 cctggcaccaccataagccagcctcgggagggcaccacctaa 345  
DB 401 cctggcaccaccataagccagcctcgggagggcaccacctaa 445

RESULT 8  
AA159150  
ID AA159150 standard; cDNA; 771 BP.  
XX  
XX AA159150;  
XX  
XX 22-OCT-2001 (first entry)  
XX  
XX Human polynucleotide SEQ ID NO 1353.  
XX  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200153312-A1.  
XX

Pd		26-JUL-2001.	
Xx			
Pf	26-DEC-2000;	2000MO-US34263.	
Xx			
Pf	21-JAN-2000;	2000US-0488725.	
Xx			
Pf	25-APR-2000;	2000US-0552317.	
Pr	09-JUL-2000;	2000US-0598042.	
Pr	19-JUL-2000;	2000US-0620312.	
Pr	03-AUG-2000;	2000US-0653450.	
Pr	14-SEP-2000;	2000US-0662191.	
Pr	19-OCT-2000;	2000US-0693036.	
Pr	29-NOV-2000;	2000US-0727344.	
Xx			
Pa	(HYSE-) HYSEQ INC.		
Pi	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
Pi	Meng J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;		
Pi	Zhao QA, Zhou P, Goodrich R, Drmanac RT;		
Xx			
Dr	WPI: 2001-442253/47.		
DR	P-FSDB; AAM39994.		
Xx			
Pt	Novel nucleic acids and polypeptides, useful for treating disorders		
Pt	such as central nervous system injuries -		
Px	Claim 1; SEQ ID NO 1353; 10078bp; English.		
Xx			
Cc	The invention relates to human nucleic acids (AAI57798-AAI61369) and		
Cc	the encoded polypeptides (AAM38642-AAM42213) with nootropic,		
Cc	immunosuppressant and cyrostatic activity. The polynucleotides are useful		
Cc	in gene therapy. A composition containing a polypeptide or polynucleotide		
Cc	of the invention may be used to treat diseases of the peripheral nerve		
Cc	system, such as peripheral nervous injuries, peripheral neuropathy and		
Cc	alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic		
Cc	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
Cc	utilisation of the activities such as: Immune system suppression,		
Cc	Activator/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
Cc	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
Cc	assays for receptor activity, arthritis and inflammation, leukemias and		
Cc	C.N.S disorders.		
Cc	Note: The sequence data for this patent did not form part of the printed		
Cc	specification.		
SQ			
	Sequence 771 BP; 221 A; 203 C; 204 G; 143 T; 0 other;		
	Query Match	100.0%; Score 345; DB 22; Length 771;	
	Best Local Similarity	100.0%; Pred. No. 2.9e-86;	
	Matches 345; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 atgactctgtaaaatgtcgagctgtgaagcgcaacatagaagaccatcatcaaaccttcac	60	
Db	118 atgactctgtaaaatgtcgagctgtgaagcgcaacatagaagaccatcatcaaaccttcac	177	
OY	61 caatactctgtaaagtgtgggcacccagacacctgaaccagggtgggaattcaaaagctg	120	
Db	178 caatactctgtaaagtgtgggcacccagacacctgaaccagggtgggaattcaaaagctg	237	
OY	121 gtgcgaaaaaatctccaattttctcaaagaagsgaataagaatgaagaaggtcataga	180	
Db	238 gtgcgaaaaaatctccaattttctcaaagaagsgaataagaatgaagaaggtcataga	297	
OY	181 cacatcatggaagaccttgagacaacaatgatgagacaagcgttagcttggagagttcacc	240	
Db	298 cacatcatggaagaccttgagacaacaatgatgagacaagcgttagcttggagagttcacc	357	
OY	241 atgcgtgatggcgaagctaaccctgggctcccacggaagaatgcacgaggttacagagc	300	
Db	358 atgcgtgatggcgaagctaaccctgggctcccacggaagaatgcacgaggttacagagc	417	
OY	301 cctggcacccccataagccagcgctctggggagggcaccccctaa	345	

```

Db      418  cctggccaccacataagccaagcctcggggaaggaccacctta 462

RESULT      9
AAFS4717/c
ID  AAF54717 standard; DNA: 452 BP.
XX
XX
XX  AAF54717;
AC
XX  15-MAY-2001 (first entry)
DT
XX
XX  Nucleotide sequence of a human polynucleotide sequence.
DE
XX
XX  Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
KW  ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
KW  neurological disease; auto-immune disease; multiple sclerosis; toxicity;
KW  Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW  rheumatoid polyarthritits; lupus erythematosus; gene therapy; ss.
XX
XX  Homo sapiens.
OS
XX
XX  MO200105422-A2.
PN
XX
XX  25-JAN-2001.
PD
XX
XX  17-JUL-2000; 2000MO-FR02057.
PF
XX
XX  15-JUL-1999; 99FR-0009372.
PR
XX
XX  (INNR ) BIOMERIEUX STELHYS.
PA
XX
XX  Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;
PI
XX  WPI: 2001-159475/16.
DR
XX
XX  Detecting, preventing and treating degenerative, neurological and
PT  autoimmune diseases, particularly multiple sclerosis, using specified
PT  polypeptides or related nucleic acid or ligand
PI
XX
XX  Claim 11; Page 186-187; 209pp; French.
PS
XX
XX  The present sequence represents a human polynucleotide sequence, which
CC  is used in the method of the invention. The specification describes a
CC  method which uses at least one polypeptide or polynucleotide sequence
CC  belonging to the perlecan, precursor of the retinol-binding plasma
CC  protein, precursor of the ganglioside GM2 activator, calgranulin B or
CC  saposin B protein families. The method is used for detecting,
CC  preventing or treating a degenerative, neurological and/or auto-immune
CC  disease. The polynucleotides and polypeptides are used for diagnosis,
CC  prognosis, prevention and treatment of multiple sclerosis (in its
CC  various forms and phases). They may also be useful in cases of
CC  e.g. Alzheimer's and Parkinson's diseases, amyotrophic lateral sclerosis,
CC  rheumatoid polyarthritits and lupus erythematosus, including use as
CC  vaccines and in gene therapy (expression of sense or antisense
CC  sequences). They can also be used to assess efficacy of potential
CC  therapeutic agents, particularly compounds that reduce or inhibit
CC  toxicity towards glial cells.
CC
XX
XX  Sequence 452 BP; 69 A; 131 C; 129 G; 123 T; 0 other;
SQ

Query Match      77.5%; Score 267.4; DB 22; Length 452;
Best Local Similarity 99.6%; Pred. No. 9,1e-65;
Matches 268; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY  77  tggggcaccccaacaccccgaaacccgagggaattccaagaagctggtgcaaaaagatctgc 136
DB  452  TGGGGCACCCCAACACCCCGAACCCTGAACACGGGGGAATTCAAAAGCTGGTGGCAAAAGATCTGC 393
DB  392  AAAATTTTCTCAAGAAAGGAGTAATAGAAAGAAAGGTCTATAGAACATCATGTGAGAGACC 333

```

QY 197 tgcacacaatgacagacagctgagctcgcagaggtcattcatatgctgagcagc 256  
 Db 332 TGGACACAAATGACAGACAGCTGAGCTTCGAGAGATTTCATCATGCTGATGGCAGAGC 273  
 QY 257 taactggtgctcccaagagaagatgcagaggtgacagagggccctggccaccacata 316  
 Db 272 TAACCTGGGCTCCACACGAAATGCACGAGGTGACGAGGGCCCTGGCCACACACATA 213  
 QY 317 agccagagctcgagagggacccccctaa 345  
 Db 212 AGCCAGGCTCGGGAGGGCACCCTTA 184

RESULT 10  
 AAF54709  
 ID AAF54709 standard; DNA; 342 BP.  
 AC AAF54709;  
 DT 15-MAY-2001 (first entry)  
 DE Nucleotide sequence of precursor of retinol-binding plasma protein.  
 XX  
 XX Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;  
 KM ganglioside GM2 activator; saposin B; degenerative disease; glial cell;  
 KM neurological disease; auto-immune disease; multiple sclerosis; toxicity;  
 KM Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
 KM rheumatoid polyarthritis; lupus erythematosus; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200105422-A2.  
 PD 25-JAN-2001.  
 DE 17-JUL-2000; 2000WO-FR02057.  
 PR 15-JUL-1999; 99FR-0009372.  
 XX  
 XX (INMR ) BIOMERIEUX STELHYS.  
 PA  
 XX  
 PI Roeklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;  
 DR WPI: 2001-159475/16.  
 XX  
 DR P-PSDB; AAB31905.  
 PT Detecting, preventing and treating degenerative, neurological and  
 PT autoimmune diseases, particularly multiple sclerosis, using specified  
 PT polypeptides or related nucleic acid or ligand -  
 XX  
 XX  
 PS Claim 11; Page 183; 209pp; French.  
 XX  
 XX The present sequence represents a human polynucleotide, which is used  
 CC in the method of the invention. The specification describes a method  
 CC which uses at least one polypeptide or polynucleotide sequence  
 CC belonging to the perlecan, precursor of the retinol-binding plasma  
 CC protein, precursor of the ganglioside GM2 activator, calgranulin B or  
 CC saposin B protein families. The method is used for detecting,  
 CC preventing or treating a degenerative, neurological and/or auto-immune  
 CC disease. The polynucleotides and polypeptides are used for diagnosis,  
 CC prognosis, prevention and treatment of multiple sclerosis (in its  
 CC various forms and phases). They may also be useful in cases of  
 CC e.g. Alzheimer's and Parkinson's diseases, amyotrophic lateral sclerosis,  
 CC rheumatoid polyarthritis and lupus erythematosus, including use as  
 CC vaccines and in gene therapy (expression of sense or antisense  
 CC sequences). They can also be used to assess efficacy of potential  
 CC therapeutic agents, particularly compounds that reduce or inhibit  
 CC toxicity towards glial cells.  
 CC  
 SQ Sequence 342 BP; 89 A; 31 C; 54 G; 39 T; 129 other;

Query Match 76.0%; Score 262.2; DB 22; Length 342;

Best Local Similarity 62.5%; Pred. No. 2, 3e-63;  
 Matches 213; Conservative 86; Mismatches 42; Indels 0; Gaps 0;

QY 1 atgacttgcacaatgtcgcagctgtgacacacatagagacatcatcaacaccttcac 60  
 Db 1 atgacttgyaartgtgncarytngarmgnaayathgaracnaththaaycnytcay 60  
 QY 61 caatactgtgaagctggggacccagacacccctgacacaggggaattcaagaagctg 120  
 Db 61 cartaywsngtnaarytnguncaycngayacnytnaaycargngarttlyaaraytn 120  
 QY 121 gtgcgaaaagatctgcacaaatttctcaagaagagaaataagaatgaaggltcataga 180  
 Db 121 gltmgnaaragayltnaarayltnaaraylaaraayaaraayaaragaytnatngar 180  
 QY 181 cacatcatgagagctgtgacacacaaatgacagacagctgagcttcgaggaagttcatc 240  
 Db 181 cayathatgagagaytngayacnaaycngayaaarcarytmsntlygargarttyath 240  
 QY 241 atgctatgagagagctaaccttgagctccacagagaagatgcagaggttgcagagggc 300  
 Db 241 atgttnatgagcmngytnacnttgagcmwncaygataaratgagayargngayargn 300  
 QY 301 cctggccacacataagccagagctcgaggagggaccccc 341  
 Db 301 ccngnncaycaycaarccmngytnngngargnagcnc 341

RESULT 11  
 AA160936  
 ID AA160936 standard; cDNA; 345 BP.  
 AC AA160936;  
 DT 22-OCT-2001 (first entry)  
 DE Human polynucleotide SEQ ID NO 4925.  
 XX  
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;  
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 PN WO200153312-A1.  
 PD 26-JUL-2001.  
 DE 26-DEC-2000; 2000WO-US34263.  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0528317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR P-PSDB; AAM41780.  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX Claim 1: SEQ ID NO 4925; 10078pp; English.  
PS  
XX  
CC The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA38642-AA42213) with nucleotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
CC  
XX  
SQ Sequence 345 BP; 110 A; 86 C; 81 G; 68 T; 0 other;

Query Match 75.8%; Score 261.6; DB 22; Length 345;  
Best Local Similarity 88.0%; Pred. No. 3.3e-63;  
Matches 285; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1 atgacttgcaaaatgtcgcagctggaacgcacacatagaccatcatcaacaccttcac 60  
Db 21 atgacttgcaaaatgtcgcagctggaacgcacacatagaccatcatcaacaccttcac 80  
QY 61 caatactctgtgaagctcggggcaccgcagacacccctgaaccagggggaattcaagaagctg 120  
Db 81 catactctgaagagctcggggcaccgcagacacccctgaaccagggggaattcaagaagctg 140  
QY 121 gtgcgaaaagaatctgcgaatatttctcaagaagagaataagaatgaaggtcatagaa 180  
Db 141 gtgcgaaagagctcgtcaataacattcctcaagaagagaataagaatgaaggtcatagaa 200  
QY 181 cacatcatggaagacccctggacacaatgacagaacagctgagcttcgagaggtcatc 240  
Db 201 cacatcatggaagacccctggacacaatgacagaacagctgagcttcgagaggtcatc 260  
QY 241 atgctgtgtgagagcttaacctggtgctccacagagaagatgacagaggtgacagaggc 300  
Db 261 atgctgtgtgagagcttaacctggtgctccacagagaagatgacagaggtgacagaggc 320  
QY 301 cctggcaccaccataaagccagagc 324  
Db 321 cctgtgcaacaacataagccagagc 344

RESULT 12  
AAC06727  
ID AAC06727 standard; cDNA; 385 BP.  
XX  
AC AAC06727;  
XX  
XX 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 10802.  
XX  
XX  
KM Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-0200610.  
XX

PR 26-FEB-1999; 99US-0122487.  
XX  
XX  
PA (GEST) GENSET.  
XX  
XX Dumas Maline Edwards J, Duclert A, Giordano J;  
XX  
XX WPI; 2000-500381/45.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
XX Claim 1: SEQ ID 10802; 71pp + CD-ROM; English.  
XX  
XX The present sequence is one of a large number of 5' ESTs derived from  
XX cDNAs encoding secreted proteins. No ORF has yet been conclusively  
XX identified within the present sequence. The 5' ESTs were prepared from  
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
XX sequences usually correspond mainly to the 3' untranslated region (UTR)  
XX of the mRNA because they are often obtained from oligo-dT primed cDNA  
XX libraries. Such ESTs are not well suited for isolating cDNA sequences  
XX derived from the 5' ends of mRNAs and even in those cases where longer  
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.  
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX They are used to obtain upstream regulatory sequences and to design  
XX expression and secretion vectors.  
XX  
SQ Sequence 385 BP; 116 A; 104 C; 104 G; 59 T; 2 other;

Query Match 75.4%; Score 260.2; DB 21; Length 385;  
Best Local Similarity 89.1%; Pred. No. 8.6e-63;  
Matches 304; Conservative 2; Mismatches 0; Indels 35; Gaps 1;

QY 1 atgacttgcaaaatgtcgcagctggaacgcacacatagagaccatcatcaacaccttcac 60  
Db 48 atgacttgcaaaatgtcgcagctggaacgcacacatagagaccatcatcaacaccttcac 107  
QY 61 caatactctgtgaagctcggggcaccgcagacacccctgaaccagggggaattcaagaagctg 120  
Db 108 catactctgtgaagctcggggcaccgcagacacccctgaaccagggggaattcaagaagctg 167  
QY 121 gtgcgaaaagaatctgcgaatatttctcaagaagagaataagaatgaaggtcatagaa 180  
Db 168 gtgcgaaaagaatctgmaaatv-----gaa 192  
QY 181 cacatcatggaagacccctggacacaatgacagaacagctgagcttcgagaggtcatc 240  
Db 193 cacatcatggaagacccctggacacaatgacagaacagctgagcttcgagaggtcatc 252  
QY 241 atgctgtgtgagagcttaacctggtgctccacagagaagatgacagaggtgacagaggc 300  
Db 253 atgctgtgtgagagcttaacctggtgctccacagagaagatgacagaggtgacagaggc 312  
QY 301 cctggcaccaccataaagccagagcctcggggagggcagccc 341  
Db 313 cctggcaccaccataaagccagagcctcggggagggcagccc 353

RESULT 13  
AAF54715/C  
ID AAF54715 standard; DNA; 430 BP.  
XX  
XX AAF54715;  
XX  
XX 15-MAY-2001 (first entry)  
XX  
XX  
DE Nucleotide sequence of a human polynucleotide sequence.  
XX  
XX  
KW Human: perlecan; retinol-binding plasma protein; calgranulin B; vacuole;  
KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;

OS	Homo sapiens.
XX	
XX	
XX	MO200105422-A2.
PB	25 -JAN-2001.
PD	
PE	17-JUL-2000: 2000MO-FR02057.
PF	
PR	15-JUL-1999: 99FR-0009372.
PS	
PA	(INMR ) BIOMERIEUX STELHYS.
XX	
PI	Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;
DR	WPI; 2001-159475/16.
XX	
PT	Detecting, preventing and treating degenerative, neurological and
PT	autoimmune diseases, particularly multiple sclerosis, using specified
PT	polypeptides or related nucleic acid or ligand
XX	
PS	Claim 11; Page 186; 209pp; French.
CC	The present sequence represents a human polynucleotide sequence, which
CC	is used in the method of the invention. The specification describes a
CC	method which uses at least one polypeptide or polynucleotide sequence
CC	belonging to the perlecan, precursor of the retinol-binding plasma
CC	protein, precursor of the ganglioside GM2 activator, calgranulin B or
CC	saposin B protein families. The method is used for detecting,
CC	preventing or treating a degenerative, neurological and/or auto-immune
CC	disease. The polynucleotides and polypeptides are used for diagnosis,
CC	prognosis, prevention and treatment of multiple sclerosis (in its
CC	various forms and phases). They may also be useful in cases of
CC	e.g. Alzheimer's and Parkinson's diseases, amyotrophic lateral sclerosis,
CC	rheumatoid polyarthritis and lupus erythematosus, including use as
CC	vaccines and in gene therapy (expression of sense or antisense
CC	sequences). They can also be used to assess efficacy of potential
CC	therapeutic agents, particularly compounds that reduce or inhibit
CC	toxicity towards glial cells.
XX	
SQ	Sequence 430 BP; 68 A; 125 C; 120 G; 117 T; 0 other:
Query Match	70.1%; Score 242; DB 22; Length 430;
Best Local Similarity	100.0%; Pred. No. 1e-57;
Matches 242; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	104 gggagattcaaaagcctggtgcgaaaagaatcycacaatlcttccaagaagaataaaga 163
DB	430 GGGAAATTCAAAAGSGCTGCGTGCAGMAAAGATCCTCAAAATTCTTCACAGAGAGATTAAGA 371
OY	164 atgaaaaagtcataagaacacatcatctgtaggaactctggaacaacaaatgcagacaagcactga 223
DB	370 ATGAAAAAGGTCAATGAACAACATCATGAGAGACTTGGAACAAATAAGTCAGAACAGACCTGA 311
OY	224 gcttcgaggagtctcatcatcgtgatcgagagctaaccttgggcctccacagagaagaatgc 283
DB	310 GCTTCGAGGAGTTTCATCATGCTGAGTAGGGAGGCTTAACCTGGGCTCCACAGAGAAGTGG 251
OY	284 acgagaggtgacgagggcctctgacacacacataagcagagcctctgggagaggaaccccct 343
DB	250 ACGAGGGGTGACGAGGGGCCCTGGCACACACACATTAAGCCAGGCTCGGGAGAGGACACCCTCT 191
OY	344 aa 345
DB	190 AA 189

RESULT 14  
AAF54713/c

ID	AAFS4713 standard; DNA: 425 BP.
AC	AAFS4713;
DT	15-MAY-2001 (first entry)
XX	
DE	Nucleotide sequence of a human polynucleotide sequence.
XX	
KW	Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
KW	ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
KW	neurological disease; auto-immune disease; multiple sclerosis; toxicity;
KW	Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW	rheumatoid polyarthritis; lupus erythematosus; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200105422-A2.
PD	25-JAN-2001.
XX	
PF	17-JUL-2000; 2000MC-FR02057.
XX	
PR	15-JUL-1999; 99FR-0009372.
XX	
PA	(INMR ) BIOMERIEUX STELHYS.
XX	
PI	Roecklin D, Kolbe H, Charles M, Maucus C, Santoro L, Perron H;
DR	WPI: 2001-159475/16.
XX	
PT	Detecting, preventing and treating degenerative, neurological and
PT	autoimmune diseases, particularly multiple sclerosis, using specified
PT	polypeptides or related nucleic acid or ligand
XX	
PS	Claim 11; Page 185-186; 209pp; French.
XX	
CC	The present sequence represents a human polynucleotide sequence, which
CC	is used in the method of the invention. The specification describes a
CC	method which uses at least one polypeptide or polynucleotide sequence
CC	belonging to the perlecan, precursor of the retinol-binding plasma
CC	protein, precursor of the ganglioside GM2 activator, calgranulin B or
CC	saposin B protein families. The method is used for detecting,
CC	preventing or treating a degenerative, neurological and/or auto-immune
CC	disease. The polynucleotides and polypeptides are used for diagnosis,
CC	prognosis, prevention and treatment of multiple sclerosis (in its
CC	various forms and phases). They may also be useful in cases of
CC	e.g. Alzheimer's and Parkinson's diseases, amyotrophic lateral sclerosis,
CC	rheumatoid polyarthritis and lupus erythematosus, including use as
CC	vaccines and in gene therapy (expression of sense or antisense
CC	sequences). They can also be used to assess efficacy of potential
CC	therapeutic agents, particularly compounds that reduce or inhibit
CC	toxicity towards glial cells.
XX	
XX	Sequence 425 BP; 68 A; 124 C; 117 G; 115 T; 1 other;
QY	Query Match 69.98; Score 241; DB 22; Length 425;
DB	Best Local Similarity 99.68; Pred. No. 2e-57;
DB	Matches 241; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY	104 ggggaattcaagagactcgtgacgaagaagatctgcaaatcttcacgaagaggaataaga 163
DB	425 GGGAAATTCAAGAGACTCGTGCGAAGAAGATCTGCAAAATTTCTCAAGAAGAGATTAAG 366
QY	164 atgaaaaggtcatagaaacacatcatgtagagactctgacacaatgcaagaacagctgtga 223
DB	365 ATGAAGAAGTCATATGAAACACATCATGTAGAGAGACCTCGACACAATGACAGCAAGCTGA 306
QY	224 gtttcgagagattatcatatgcttagtgcgagggcttaaccctcggagcctccacggaagatgc 283
DB	305 GCTTCGAGAGATTATCATCTAGTCTAGTGGCGAGGCTTAACCTGGGCGCTCCACAGAAAGATGC 246
QY	284 acgagagtgacgagagccctctggccacacacataagccaggcctcggggaaggccacctt 343

Db 245 ACCAGGGTGCAGGAGGCGCTGGCCACCACCATAGCCAGGCGCTGGGAGGAGCCCTT 186  
QY 344 aa 345  
11  
Db 185 AA 184

RESULT 15  
ABA59810/C  
ID ABA59810 standard: DNA; 592 BP.

XX ABA59810:

AC ABA59810:

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #8115.

XX Human: foetal liver: gene expression: single exon nucleic acid probe: ss.

XX Homo sapiens.

XX MO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001MO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human foetal liver -

XX Claim 1; SEQ ID NO 8115; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

XX measuring human gene expression in a sample derived from human foetal

XX liver. The single exon nucleic acid probes may be used for predicting,

XX measuring and displaying gene expression in samples derived from human

XX foetal liver. The present sequence is a single exon nucleic acid

XX probe of the invention.

XX Note: The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 592 BP; 111 A; 150 C; 191 G; 140 T; 0 other;

Query Match 57.1%; Score 197; DB 22; Length 592;

Best Local Similarity 100.0%; Pred. No. 4e-45;

Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 agaaggaataaagatgaagagcatagacacatcatgtagagagcctgagacacaatg 208

Db 526 AGAAGGAAATTAAGATGAAGGTCATAGAACATCATGAGAGCTGGACACAATG 467

QY 209 cagacaagcagctgagctcgagaggttcacatgctgagcgaggttaacctggcct 268

Db 466 CAGACAAGCAGCTGAGCTTCGAGGATTCATCATGCTGATGGAGGCTAACCTGGGCT 407

QY 269 Tccatgagagaatgcaagaggtgaagagggccttgccaccacataaagccagcctcg 328

Db 406 CCCACGAGAAAGATGCAAGAGGAGGAGGCGCTGGCCACACCATTAAGCAAGGCTCG 347  
QY 329 gggagggcaccacctaa 345  
11  
Db 346 GGGAGGGCACCCCTTAA 330

Search completed: September 10, 2002, 06:49:30  
Job time: 5206 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2002, 06:07:54 ; Search time 2463.79 Seconds  
(without alignments)  
1889.954 Million cell updates/sec

Title: US-09-806-382a-2  
Perfect score: 345  
Sequence: 1 atgactgcgaataatgcga.....tcgaggagggcacccctaa 345

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues  
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estmu:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estcpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	345	100.0	539	10	BE551799
2	345	100.0	567	9	AI952029 wx46a07.x
3	345	100.0	582	10	BM008345 603617323
4	345	100.0	609	10	BC438100 602490368
5	345	100.0	638	9	AW964579 EST376652
6	345	100.0	754	10	BC331778 602432209
7	345	100.0	781	10	BE787206 601477976
8	345	100.0	935	10	BI907927 603069064
9	344	99.7	524	9	AA587142 mn70d09.s
10	344	99.7	786	10	BC488718 602534663
11	343.4	99.5	553	10	BE550359 ht75a03.x
12	343.4	99.5	680	10	BI837657 603086883
13	341.8	99.1	561	10	BF511375 UT-H-B14
14	341	98.8	1190	10	BE785947 601477976
15	333	96.5	500	10	BI518253 603042106
16	333	96.5	573	10	BI760899 603043772
17	333	96.5	576	10	BE786188 601474308

18	333	96.5	725	10	BC332100	BC332100 602432509
19	333	96.5	801	10	BC437703	BC437703 602488965
20	333	96.5	887	10	BC331032	BC331032 602431694
21	333	96.5	907	10	BI836342	BI836342 603083037
22	333	96.5	968	10	BI910664	BI910664 603070429
23	332.2	96.3	399	10	BI031826	BI031826 IL5-MT026
24	329.8	95.6	404	10	BE941564	BE941564 na104b01.
25	329.2	95.4	964	10	BE784331	BE784331 601473479
26	327.2	94.8	569	9	AW269689	AW269689 xv55d03.x
27	326.6	94.7	423	9	AA587152	AA587152 nn74a10.s
28	323.4	93.6	525	9	AA846123	AA846123 ak83h08.s
29	323.2	93.6	603	10	BC438172	BC438172 602490260
30	321.6	93.2	513	10	BE350415	BE350415 ht85d06.x
31	321	93.0	576	10	BE183248	BE183248 601809954
32	321	93.0	863	10	BE787379	BE787379 601479508
33	321	93.0	1015	10	BI517687	BI517687 603042106
34	320	92.8	557	9	AA586741	AA586741 nn71a07.s
35	316.4	91.7	781	10	BE240407	BE240407 601905958
36	310.4	90.0	548	9	AA583988	AA583988 nn60h02.s
37	307	89.0	547	9	AI281728	AI281728 ct73d01.x
38	303.8	88.1	437	10	W76577	W76577 zd66a09.r1
39	302	87.5	871	10	BE787611	BE787611 601481715
40	301	87.2	891	10	BI839178	BI839178 603085908
41	293	84.9	508	9	AA582879	AA582879 nn72c01.s
42	286.4	83.0	295	10	BE714495	BE714495 PM4-HT072
43	283.6	82.2	540	9	AA586920	AA586920 nn68f07.s
44	281.2	81.5	566	9	AW468067	AW468067 he31e04.x
45	279.4	81.0	307	9	AA320692	AA320692 EST23134

## ALIGNMENTS

RESULT 1  
LOCUS BE551799 539 bp mRNA linear EST 10-AUG-2000  
DEFINITION ht86d01.x1 NCI CGAP HN15 Homo sapiens CDNA clone IMAGE:3159361  
Similar to gb:X06233 CALGRANULIN B (HUMAN);, mRNA sequence.

ACCESSION BE551799  
VERSION BE551799.1 GI:9793491  
KEYWORDS EST.

SOURCE human.  
ORANISM human.

REFERENCE 1 (bases 1 to 539)  
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.

AUTHORS NCI/NIDR-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

JOURNAL National Cancer Institute / National Institute of Dental Research,  
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

Tissue Procurement: Edward J. Shillitoe, Ph.D., Chlehanok  
Leeethanakul, D.D.S.  
CDNA Library Preparation: Krizman and Leeethanakul Laboratories

DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:  
info@image.lnl.gov

Seq primer: -40up from Gidco  
High quality sequence stop: 437.  
Location/Qualifiers  
1. 539  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3159361"  
/clone\_lib="NCI CGAP HN15"  
/tissue\_type="Leukoplakia of the buccal mucosa"  
/lab\_host="DH10B"  
/note="Organ: mouth; Vector: pAMP10; mRNA made from  
leukoplakia, CDNA made by oligo-dT priming.  
Non-directionally cloned into UDG sites. Size-selected on

## FEATURES

source

agarose gel, average insert size 500 bp. Primary library.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
REFERENCE: Krizman et al. (1996) Cancer Research  
56:5380-5383 "

BASE COUNT    154 a    159 c    146 g    80 t

ORIGIN

Query Match                    100.0%; Score 345; DB 10; Length 539;  
Best Local Similarity 100.0%; Pred. No. 5.9e-71;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY    1 atgacttcgaaatgtgcagctggaagcgaactagagaccatccacacaccttcac 60  
      |||||||  
Db    62 atgacttcgaaatgtgcagctggaagcgaactagagaccatccacacaccttcac 121  
      |||||||  
QY    61 caatactctgtgaagcttggcaccgacacacccctgaaccagggggaattcaagaagctg 120  
      |||||||  
Db    122 caatgactctgtgaagcttggcaccgacacacccctgaaccagggggaattcaagaagctg 181  
      |||||||  
QY    121 gtgcgaaaaagatctgcacaaatttctcagaagaagagaataagatgaaaggtcatagaa 180  
      |||||||  
Db    182 gtgcgaaaaagatctgcacaaatttctcagaagaagagaataagatgaaaggtcatagaa 241  
      |||||||  
QY    181 cacatcatggaagacctgtgacacaaatgcagacaagcagctgagcttcgaggaagttcatc 240  
      |||||||  
Db    242 cacatcatggaagacctgtgacacaaatgcagacaagcagctgagcttcgaggaagttcatc 301  
      |||||||  
QY    241 atgctgatatggagagcttaacctgtggctcccaagagaagatgcagaggggtgcagagggc 300  
      |||||||  
Db    302 atgctgatatggagagcttaacctgtggctcccaagagaagatgcagaggggtgcagagggc 361  
      |||||||  
QY    301 cctggcaccacacataagccagcctcgaggaggagggccccctaa 345  
      |||||||  
Db    362 cctggcaccacacataagccagcctcgaggaggagggccccctaa 406  
      |||||||

RESULT    2  
AI952029    567 bp    mRNA    linear    EST 09-MAR-2000  
LOCUS    wx46a07.x1 NCI CGAP Lu28 Homo sapiens cDNA clone IMAGE:2546676 3'  
DEFINITION    similar to gb:X06233 CALGRANULIN B (HUMAN);, mRNA sequence.  
ACCESSION    AI952029  
VERSION    AI952029.1 GI:5744339  
KEYWORDS    EST.  
SOURCE    human.  
ORGANISM    Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE    1 (bases 1 to 567)  
AUTHORS    NCI-CGAP htcp://www.ncbi.nlm.nih.gov/ncicgap.  
TITLES    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
          Tumor Gene Index  
JOURNAL    Unpublished (1997)  
COMMENT    Contact: Robert Strausberg, Ph.D.  
          Email: cgapbs-remail.nih.gov  
          Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.  
          Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life  
          Technologies, Inc. CDNA Library Arrayed by: Christa Prange. The  
          I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
          Genome Sequencing Center  
          Clone distribution: NCI-CGAP clone distribution information can be  
          found through the I.M.A.G.E. Consortium/LLNL at:  
          www.bio.llnl.gov/bdrip/image/image.html  
          Insert Length: 413    Std Error: 0.00  
          Seq primer: ~400p from Gldco  
          High quality sequence stop: 408.  
          Location/Qualifiers  
          1..567  
          /organism="Homo sapiens"  
          /db\_xref="taxon:9606"  
          /clone="IMAGE:2546676"  
          /clone\_lid="NCI\_CGAP\_Lu28"

## FEATURES

source  
1..567  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2546676"  
/clone\_lid="NCI\_CGAP\_Lu28"

/tissue\_type="two pooled squamous cell carcinomas"  
/lab\_host="DH10B"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies."

BASE COUNT    91 a    157 c    159 g    160 t

ORIGIN

Query Match                    100.0%; Score 345; DB 9; Length 567;  
Best Local Similarity 100.0%; Pred. No. 5.9e-71;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY    1 atgacttcgaaatgtgcagctggaagcgaactagagaccatccacacaccttcac 60  
      |||||||  
Db    537 atgacttcgaaatgtgcagctggaagcgaactagagaccatccacacaccttcac 478  
      |||||||  
QY    61 caatactctgtgaagcttggcaccgacacacccctgaaccagggggaattcaagaagctg 120  
      |||||||  
Db    477 caatgactctgtgaagcttggcaccgacacacccctgaaccagggggaattcaagaagctg 418  
      |||||||  
QY    121 gtgcgaaaaagatctgcacaaatttctcagaagaagagaataagatgaaaggtcatagaa 180  
      |||||||  
Db    417 gtgcgaaaaagatctgcacaaatttctcagaagaagagaataagatgaaaggtcatagaa 358  
      |||||||  
QY    181 cacatcatggaagacctgtgacacaaatgcagacaagcagctgagcttcgaggaagttcatc 240  
      |||||||  
Db    357 cacatcatggaagacctgtgacacaaatgcagacaagcagctgagcttcgaggaagttcatc 298  
      |||||||  
QY    241 atgctgatatggagagcttaacctgtggctcccaagagaagatgcagaggggtgcagagggc 300  
      |||||||  
Db    297 atgctgatatggagagcttaacctgtggctcccaagagaagatgcagaggggtgcagagggc 238  
      |||||||  
QY    301 cctggcaccacacataagccagcctcgaggaggagggccccctaa 345  
      |||||||  
Db    237 cctggcaccacacataagccagcctcgaggaggagggccccctaa 193  
      |||||||

RESULT    3  
BM008345    582 bp    mRNA    linear    EST 30-OCT-2001  
LOCUS    603617323P1 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:5450668 5',  
DEFINITION    mRNA sequence.  
ACCESSION    BM008345  
VERSION    BM008345.1 GI:16522699  
KEYWORDS    EST.  
SOURCE    human.  
ORGANISM    Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE    1 (bases 1 to 582)  
AUTHORS    NIH-MGC htcp://mgc.ncbi.nlm.nih.gov/  
TITLES    National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL    Unpublished (1999)  
COMMENT    Contact: Robert Strausberg, Ph.D.  
          Email: cgapbs-remail.nih.gov  
          Tissue Procurement: Dr. Mark Watson  
          CDNA Library Preparation: Ling Hong/Rubin Laboratory  
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
          DNA Sequencing by: Incyte Genomics, Inc.  
          Clone distribution: MGC clone distribution information can be  
          found through the I.M.A.G.E. Consortium/LLNL at:  
          htcp://llnl.gov  
          Plate: LCM1939 row: m column: 05  
          High quality sequence stop: 581.  
          Location/Qualifiers  
          1..582  
          /organism="Homo sapiens"  
          /db\_xref="taxon:9606"  
          /clone="IMAGE:5450668"  
          /clone\_lid="NIH\_MGC\_113"  
          /lab\_host="DH10B (phage-resistant)"  
          /note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2:

## FEATURES

source  
1..582  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5450668"  
/clone\_lid="NIH\_MGC\_113"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2:

ECORI: cDNA made by oligo-dT priming. Directionally cloned into EcorI/XhoI sites using the following 5' adaptor: GGCGAGG(5). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 174 a 159 c 157 g 92 t  
ORIGIN

Query Match 100.0%; Score 345; DB 10; Length 582;  
Best Local Similarity 100.0%; Pred. No. 6e-71;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgactgtcaaaatgtcgacgtgtgaacgacaatagagaccatcataacaccttcac 60  
|||||  
DB 32 ATGACTTGCAAAATGTGCGAGCTGGAGACCAATAGAGACCATCATCAACCTTCAC 91  
QY 61 caatactctgtgaagctgtgggagaccagacacccctgaaccaggggaattcaagaagctg 120  
|||||  
DB 92 CAATACTCTGTGAAGCTGGGACCCAGACACCTCTGAACAGGGGGAATTCAAGAGCTG 151  
QY 121 gtgcgaaaagatctgcgaaaattttctcaagaagaagaatgaagaatgaagaatgaaga 180  
|||||  
DB 152 GTGCGAAAAGATCTGCAAAATTTTCTCAAGAGAGAAATGAAGATGAAGAGCTCATAGAA 211  
QY 181 cacatcatgtgaagacctgtgaacacaatgtgaagacagcgtgaagttcgaagaattcattc 240  
|||||  
DB 212 CACATCATGTGAGAGACCTGTGACACAAATGCAACAGCACTGCTTCAGAGAGTTTCATC 271  
QY 241 atgcgtatgtgcagagctgaacctgtggcctcccaagagaagaatgcagaggtgtgaagaaggc 300  
|||||  
DB 272 ATGCTGATGTGGCAGGCTAATCTGGGCTCCCAAGAGAAATGACAGAGGCTGACGAGGCG 331  
QY 301 cctggccacacacataagccagagcctcggggaggggacccccctaa 345  
|||||  
DB 332 CCGTGGCCACACCATTAAGCCAGGCTCGGGGAGGGGACCCCTTA 376

RESULT 4  
BG438100 609 bp mRNA linear EST 14-MAR-2001  
LOCUS 602490368F1 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:4622272 5',  
DEFINITION mRNA sequence.  
ACCESSION BG438100  
VERSION BG438100.1 GI:13344606  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 609)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: DCTD/DPH/Gazdar  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
plate: LICM1386 row: h column: 17  
High quality sequence stop: 595.  
Location/Qualifiers

FEATURES  
source 1..609  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4622272"  
/clone\_1tb="NIH\_MGC\_18"

BASE COUNT 206 a 160 c 154 g 89 t  
ORIGIN

Query Match 100.0%; Score 345; DB 10; Length 609;  
Best Local Similarity 100.0%; Pred. No. 6.1e-71;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgactgtcaaaatgtcgacgtgtgaacgacaatagagaccatcataacaccttcac 60  
|||||  
DB 28 ATGACTTGCAAAATGTGCGAGCTGGAGACCAATAGAGACCATCATCAACCTTCAC 87  
QY 61 caatactctgtgaagctgtgggagaccagacacccctgaaccaggggaattcaagaagctg 120  
|||||  
DB 88 CAATACTCTGTGAAGCTGGGACCCAGACACCTCTGAACAGGGGGAATTCAAGAGCTG 147  
QY 121 gtgcgaaaagatctgcgaaaattttctcaagaagaagaatgaagaatgaagaatgaaga 180  
|||||  
DB 148 GTGCGAAAAGATCTGCAAAATTTTCTCAAGAGAGAAATGAAGATGAAGAGCTCATAGAA 207  
QY 181 cacatcatgtgaagacctgtgaacacaatgtgaagacagcgtgaagttcgaagaattcattc 240  
|||||  
DB 208 CACATCATGTGAGAGACCTGTGACACAAATGCAACAGCACTGCTTCAGAGAGTTTCATC 267  
QY 241 atgcgtatgtgcagagctgaacctgtggcctcccaagagaagaatgcagaggtgtgaagaaggc 300  
|||||  
DB 268 ATGCTGATGTGGCAGGCTAATCTGGGCTCCCAAGAGAAATGACAGAGGCTGACGAGGCG 327  
QY 301 cctggccacacacataagccagagcctcggggaggggacccccctaa 345  
|||||  
DB 328 CCGTGGCCACACCATTAAGCCAGGCTCGGGGAGGGGACCCCTTA 372

RESULT 5  
AW964579 638 bp mRNA linear EST 01-JUN-2000  
LOCUS AW964579  
DEFINITION EST376652 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW964579  
VERSION AW964579.1 GI:8154415  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 638)  
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspar,R., Gay,C., Holt  
,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
Quackenbush,J.  
TITLE Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
JOURNAL Unpublished (2000)  
COMMENT Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: [johnq@ligr.org](mailto:johnq@ligr.org)  
Plate: 203  
Seq primer: Reverse.  
Location/Qualifiers

FEATURES  
source 1..638  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

/clone\_lib="MAGE resequences", MACH"  
/note="Vector: pbluescriptskm"  
BASE COUNT 185 a 174 c 167 g 110 t 2 others  
ORIGIN

Query Match 100.0%; Score 345; DB 9; Length 638;  
Best Local Similarity 100.0%; Pred. No. 6.2e-71;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgacttgcaaaatgctgcagctggaacgaacatagaccatcaccacacccctccac 60  
|||||  
DB 22 ATGACTTGCAAAATGCTGCAGCTGGAACGCAACATAGACCATCATCACCTTCCAC 81  
|||||  
OY 61 caatactctgtgaagctgtgggacccaagacacccctgaacccaagggaattcaaaagctg 120  
|||||  
DB 82 CAATACTCTGTGAAGCTGTGGGGACCCAGACACCTGAACCAAGGGAATCAAGAGACTG 141  
|||||  
OY 121 gtgcgaaaaagatctgcaaaatttctcaagaagaagaataagaatgaagaagtcataagaa 180  
|||||  
DB 142 GTGCGAAAAGATCTCAAAATTTTCTCAAGAAAGAAATGAATGAAGGTATAGAA 201  
|||||  
OY 181 cacatcatgaggagacctggaacaaatgcaagacagcgtgagcttcgagagatcacc 240  
|||||  
DB 202 CACATCATGAGAGACCTGGACACAATGACAGACAGCTGAGCTTCGAGAGATTTCATC 261  
|||||  
OY 241 atgctgtagtggcagagcttaacctgtggcctcccaagagaatgacagaggtgtgacagagc 300  
|||||  
DB 262 ATGCTGATGGCGAGGCTTAACCTGGGCTCCACAGAGAGATGACAGAGGTGACGAGGGC 321  
|||||  
OY 301 cctggcacaacacataaagccagcgtcgtgggagagagacccctaa 345  
|||||  
DB 322 CCTGGCACACCATTAAGCAGAGGCTCGGGAGGAGCACCCCTTA 366  
|||||

RESULT 6  
BG331778 754 bp mRNA linear EST 27-FEB-2001  
LOCUS 602432209F1 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:4549898 5',  
DEFINITION mRNA sequence.  
ACCESSION BG331778  
VERSION BG331778.1 GI:1318216  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 754)  
NIH-MGC <http://mhc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: DCPD/DTF/Gazdar  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM1242 row: a column: 03  
High quality sequence stop: 563.  
Location/Qualifiers  
1..754  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="4549898"  
/clone\_lib="NIH\_MGC\_18"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pOT87; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="4549898"  
/clone\_lib="NIH\_MGC\_18"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pOT87; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:

BASE COUNT 229 a 243 c 189 g 93 t  
ORIGIN

Query Match 100.0%; Score 345; DB 10; Length 754;  
Best Local Similarity 100.0%; Pred. No. 6.5e-71;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgacttgcaaaatgctgcagctggaacgaacatagaccatcaccacccctccac 60  
|||||  
DB 31 ATGACTTGCAAAATGCTGCAGCTGGAACGCAACATAGACCATCATCACCTTCCAC 90  
|||||  
OY 61 caatactctgtgaagctgtgggacccaagacacccctgaacccaagggaattcaaaagctg 120  
|||||  
DB 91 CAATACTCTGTGAAGCTGTGGGACCCAGACACCTGAAACAGGGGGAATTCAAGAGACTG 150  
|||||  
OY 121 gtgcgaaaaagatctgcaaaatttctcaagaagaagaataagaatgaagaagtcataagaa 180  
|||||  
DB 151 GTGCGAAAAGATCTCAAAATTTTCTCAAGAAAGAAATGAATGAAGGTATAGAA 210  
|||||  
OY 181 cacatcatgaggagacctggaacaaatgcaagacagcgtgagcttcgagagatcacc 240  
|||||  
DB 211 CACATCATGAGAGACCTGGACACAATGACAGACAGCTGAGCTTCGAGAGATTTCATC 270  
|||||  
OY 241 atgctgtagtggcagagcttaacctgtggcctcccaagagaatgacagaggtgtgacagagc 300  
|||||  
DB 271 ATGCTGATGGCGAGGCTTAACCTGGGCTCCACAGAGAGATGACAGAGGTGACGAGGGC 330  
|||||  
OY 301 cctggcacaacacataaagccagcgtcgtgggagagagacccctaa 345  
|||||  
DB 331 CCTGGCACACCATTAAGCAGAGGCTCGGGAGGAGCACCCCTTA 375  
|||||

RESULT 7  
BE787206 781 bp mRNA linear EST 20-OCT-2000  
LOCUS 601476885F1 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:3879669 5',  
DEFINITION mRNA sequence.  
ACCESSION BE787206  
VERSION BE787206.1 GI:10208404  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 781)  
NIH-MGC <http://mhc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: DCPD/DTF/Gazdar  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM9645 row: j column: 22  
High quality sequence stop: 545.  
Location/Qualifiers  
1..781  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="3879669"  
/clone\_lib="NIH\_MGC\_68"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="3879669"  
/clone\_lib="NIH\_MGC\_68"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"

/note="Organ: Lung; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT 261 a 214 c 220 g 86 t

Query Match 100.0%; Score 345; DB 10; Length 781;  
Best Local Similarity 100.0%; Pred. No. 6.5e-71;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgacttgcgaatctgcagctggaagcagcaacatagaccatcacatcaacattccac 60  
DB 10 ATGACTTGCgAAATCTCCAGCTGGAAGCAACATAGACATCATCATCAACCTTCCAC 69  
QY 61 caatactctgtgaagctgtgggacccagacacccctgaagcagggtgaattcaaaagctg 120  
DB 70 CAATACCTCTGTGAAGCTGGGGACCCAGACACCTGAACCGGGGGAATTCAAAGAGCTG 129  
QY 121 gtgcgaaaaagatctgcgaatcttctcaagaagagagaataagaaagttcatagaa 180  
DB 130 GTGCGAAAAAGATCTGCgAAATTTTCTCAAGAAAGGAAATAGAAAGGTATAGAA 189  
QY 181 caccatcatgtgaagacctgtgacacaatgtcagacagctgagcttcgagagttcatc 240  
DB 190 CACATCATGTGAAGAGCCTGTGAGACAAATGCAGAAAGCAGCTTGAGAGAGTTGATC 249  
QY 241 atgctgatatgtgagagcttaacctgtggcctccacagagaagatgtgcaggggtgaagggc 300  
DB 250 ATGCTGATGTGGCAGAGCTTAACCTGGGCTCTCCACGAGAAAGATGCAGAGGTGACGAGGGC 309  
QY 301 ccttgccaccacataagccagcctcggggagggcagcccccctaa 345  
DB 310 CCGGGCACACCAATAGCAGAGCCTCGGGAGGGCACCCCTTAA 354  
RESULT 8  
BI907927 935 bp mRNA linear EST 16-OCT-2001  
LOCUS 603069064F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5218046 5',  
DEFINITION mRNA sequence.  
ACCESSION BI907927  
VERSION BI907927.1 GI:16170822  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 935)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LLM11547 row: p column: 15  
High quality sequence stop: 574.  
Location/Qualifiers  
1..935  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5218046"  
/clone\_lib="NIH\_MGC\_118"  
/cise\_type="Leukocyte"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV

(destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH\_MGC Library."

BASE COUNT 309 a 281 c 248 g 97 t

Query Match 100.0%; Score 345; DB 10; Length 935;  
Best Local Similarity 100.0%; Pred. No. 6.9e-71;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgacttgcgaatctgcagctggaagcagcaacatagaccatcacatcaacattccac 60  
DB 31 ATGACTTGCgAAATGTCCAGCTGGAAGCAACATAGAGACATCATCAACCTTCCAC 90  
QY 61 caatactctgtgaagctgtgggacccagacacccctgaagcagggtgaattcaaaagctg 120  
DB 91 CAATACCTCTGTGAAGCTGGGGACCCAGACACCTGAACGAGGGGAATTCAAAGAGCTG 150  
QY 121 gtgcgaaaaagatctgcgaatcttctcaagaagagagaataagaaagttcatagaa 180  
DB 151 GTGCGAAAAAGATCTGCgAAATTTTCTCAAGAAAGGAAATAGAAATGAAAGGTCATAGAA 210  
QY 181 caccatcatgtgaagacctgtgacacaatgtcagacagctgagcttcgagagttcatc 240  
DB 211 CACATCATGTGAAGAGCCTGTGAGACAAATGCAGAAAGCAGCTTGAGAGAGTTGATC 270  
QY 241 atgctgatatgtgagagcttaacctgtggcctccacagagaagatgtgcaggggtgaagggc 300  
DB 271 ATGCTGATGTGGCAGAGCTTAACCTGGGCTCTCCACGAGAAAGATGCAGAGGTGACGAGGGC 330  
QY 301 ccttgccaccacataagccagcctcggggagggcagcccccctaa 345  
DB 331 CCGGGCACACCAATAGCAGAGCCTCGGGAGGGCACCCCTTAA 375  
RESULT 9  
AA587142 524 bp mRNA linear EST 26-SEP-1997  
LOCUS n70d009.s1 NCI CGAP Lab1 Homo sapiens cDNA clone IMAGE:1089233 3',  
DEFINITION similar to gb:X06233 CALGRANDLIN B (HUMAN);, mRNA sequence.  
ACCESSION AA587142  
VERSION AA587142.1 GI:2397956  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 524)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Stratagene, Inc.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
www-bio.lnl.gov/bbrp/image/image.html  
Insert Length: 650 Std Error: 0.00  
Seq primer: ~40m3 fwd. ET from Amersham  
High quality sequence stop: 322.  
Location/Qualifiers  
1..524  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1089233"  
/clone\_lib="NCI-CGAP\_Lar1"  
/tissue\_type="Larynx"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: larynx; Vector: Bluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGCGACGAG  
3' 3' adaptor sequence: 0.9 kb." Average insert size: 0.9 kb."  
BASE COUNT 83 a 145 c 148 g 147 t 1 others  
ORIGIN

Query Match 99.7%; Score 344; DB 9; Length 524;  
Best Local Similarity 99.7%; Pred. No. 1e-70; 1; Indels 0; Gaps 0;  
Matches 344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 atgacttgcaaaatgcgagctggaagcaacataagagaccatcacaacctccac 60  
DB 500 ATGACTTGCAAAATGTCGAGCTGGAGCAGCAACATAGACATCATCAACCTTCCAC 441  
QY 61 caatactctgtgaagctggaagcaccacacacacacacacacacacacacacac 120  
DB 440 CAATACCTGTGAGCTGGAGCAGCAGCAACCTGAAACGAGGGAATTCAGAGAGCTG 381  
QY 121 gtgcgaaaagatctgcaaaatttctcaagaagaagaataagaatgaaggtcagaa 180  
DB 380 GTGCGAAAAGATCTGCAAAATTTCTCAAGAGAGAAATTAAGAAAGTATAGAA 321  
QY 181 cacatcatggaagacctggaacacaatgcaagaagcagctgagctcgagagttcatc 240  
DB 320 CACATCATGAGAGACTGGACACAATGACAGCAAGAGCTGAGACTTCGAGAGATTATC 261  
QY 241 atgcgtatgagcgaagctaaacctgggctccacagagaagaatgacagaggtgacgaagc 300  
DB 260 ATGCTGATGGCAGAGCTAACTGGGCTCCACAGAGAGATGACAGAGGTGACGAGAGGC 201  
QY 301 cctggcaccacacataaagcagcctcgaggagagggacacccctaa 345  
DB 200 CCTGGCCACACCATTAAGCAGGCTCGGGGAGGCGACCCCTTAA 156

RESULT 10  
BG488718 786 bp mRNA linear EST 27-MAR-2001  
LOCUS 602534663F1 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:4662351 5',  
DEFINITION mRNA sequence.  
ACCESSION BG488718  
VERSION BG488718.1 GI:13450225  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 786)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: DCPD/DRP/Gazdar  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>

FEATURES  
source 1..786  
location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4662351"  
/clone\_lib="NIH-MGC\_18"  
/tissue\_type="large cell carcinoma"  
/lab\_host="Dh10B (phage-resistant)"  
/note="Organ: lung; Vector: pORF7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(C). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH-MGC Library."  
BASE COUNT 268 a 177 c 214 g 126 t 1 others  
ORIGIN

Query Match 99.7%; Score 344; DB 10; Length 786;  
Best Local Similarity 99.7%; Pred. No. 1.1e-70; 1; Indels 0; Gaps 0;  
Matches 344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 atgacttgcaaaatgctgcagctggaagcacaacataagaacacatcacaacctccac 60  
DB 27 ATGACTTGCAAAATGTCGAGCTGGAGCAGCAACATAGACATCATCAACCTTCCAC 86  
QY 61 caatactctgtgaagctggaagcaccacacacacacacacacacacacacacac 120  
DB 87 CAATACCTGTGAGCTGGAGCAGCAGCAACCTGAAACGAGGGAATTCAGAGAGCTG 146  
QY 121 gtgcgaaaagatctgcaaaatttctcaagaagaagaataagaatgaaggtcagaa 180  
DB 147 GTGCGAAAAGATCTGCAAAATTTCTCAAGAGAGAAATTAAGAAAGTATAGAA 206  
QY 181 cacatcatggaagacctggaacacaatgcaagaagcagctgagctcgagagttcatc 240  
DB 207 CACATCATGAGAGACTGGACACAATGACAGCAAGAGCTGAGCTTCGAGAGATTATC 266  
QY 241 atgcgtatgagcgaagctaaacctgggctccacagagaagaatgacagaggtgacgaagc 300  
DB 267 ATGCTGATGGCAGAGCTAACTGGGCTCCACAGAGAGATGACAGAGGTGACGAGAGGC 326  
QY 301 cctggcaccacacataaagcagcctcgaggagagggacacccctaa 345  
DB 327 CCTGGCCACACCATTAAGCAGGCTCGGGGAGGCGACCCCTTAA 371

RESULT 11  
BE350359 553 bp mRNA linear EST 18-JUL-2000  
LOCUS ht75a03.x1 NCI-CGAP\_HN15 Homo sapiens cDNA clone IMAGE:3158284  
DEFINITION Similar to gb:X06253 CALGRANULIN B (HUMAN);, mRNA sequence.  
ACCESSION BE350359  
VERSION BE350359.1 GI:9262212  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 553)  
AUTHORS NCI/NIH-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute / National Institute of Dental Research,  
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Edward J. Shillito, Ph.D., Chidehanok  
Leechankul, D.D.S.  
CDNA Library Preparation: Krizman and Leethanankul Laboratories  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:

info@image.llnl.gov  
Seq primer: -40UP from Gldpc  
High quality sequence stop: 463.  
Location/Qualifiers

## FEATURES

source

1..553  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3158284"  
/clone\_lib="NCI CGAP HN15"  
/tissue\_type="Leukoplakia of the buccal mucosa"  
/lab\_host="DH10B"  
/note="Organ: mouth; Vector: pAMP10; mRNA made from leukoplakia, cDNA made by oligo-dt priming. Non-directionally cloned into UDG sites. Size-selected on agarose gel, average insert size 500 bp. Primary library. cDNA library Preparation: David B. Krizman, Ph.D. REFERENCE: Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 153 a 164 c 149 g 87 t  
ORIGIN

Query Match 99.5%; Score 343.4; DB 10; Length 553;  
Best Local Similarity 99.7%; Pred. No. 1.4e-70;  
Matches 344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgacttgcaaaatgtgcagctggaacgcaacatagagacatcatcaaccttcac 60  
|||||  
DB 70 ATGACTTGCAAAATGTGCCAGCTGGACGCAACATAGACCATCATCAACCTTCAC 129  
|||||  
QY 61 caatactctgtgaagctggtgggacccagacacctgaaccagggggaattcaagaagctg 120  
|||||  
DB 130 CAATACTCTGTGAAGCTGTGGGACCCAGACACCTTGAACCGAGGGGAATTCAAAAGAGCTG 189  
|||||  
QY 121 gtgcgaaaagatctcgaattttctcaagaaggagaaataagataagaaagtcatagaa 180  
|||||  
DB 190 GTGCGAAAAGACTCGAAATTTTCTCAAGAGGAGAAATAGAAATGAAGGTCTATAGAA 249  
|||||  
QY 181 cacatcatggaagacctgtgacacaatgcagaacagcagctgagcttcgaggaattcaltc 240  
|||||  
DB 250 CACATCATGAGAGACCTGTGACACAATGACAGACAGCTGAGCTTGGAGAGATTTCATTC 309  
|||||  
QY 241 atgtctgatggaagcttaacctgtggcctcccaagagaagatgacagaggtgagcagaggtc 300  
|||||  
DB 310 ATGCTGATGGCAGAGGCTAAGCTGTGGCTCCCAAGAGAGATGCACAGAGGTGACGAGGAC 369  
|||||  
QY 301 cctggccaccacataagcagcctcgaggagagggcaccctctaa 345  
|||||  
DB 370 CCTGGCCACCACATTAAGCCAGGCTCGGGGAGGGACCCCTTAA 414  
|||||

RESULT 12  
BI837657 680 bp mRNA linear EST 04-OCT-2001  
LOCUS 603086883F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5225853 5',  
DEFINITION mRNA sequence.  
ACCESSION BI837657  
VERSION BI837657.1 GI:15949207  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 680)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM1568 row: e column: 22  
High quality sequence stop: 562.  
Location/Qualifiers

## FEATURES

source

1..680  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5225853"  
/clone\_lib="NIH\_MGC\_120"  
/lab\_host="DH10B"  
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPOK6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dt primed and directionally cloned (PCR) site is destroyed upon cloning). Average insert size 1.5 kb; insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH-MGC Library."

BASE COUNT 209 a 190 c 185 g 96 t  
ORIGIN

Query Match 99.5%; Score 343.4; DB 10; Length 680;  
Best Local Similarity 99.7%; Pred. No. 1.5e-70;  
Matches 344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgacttgcaaaatgtgcagctggaacgcaacatagagacatcatcaacaccttcac 60  
|||||  
DB 29 ATGACTTGCAAAATGTGCCAGCTGGACGCAACATAGACCATCATCAACCTTCAC 88  
|||||  
QY 61 caatactctgtgaagctggtgggacccagacacctgaaccagggggaattcaagaagctg 120  
|||||  
DB 89 CAATACTCTGTGAAGCTGTGGGACCCAGACACCTTGAACCGAGGGGAATTCAAAAGAGCTG 148  
|||||  
QY 121 gtgcgaaaagatctcgaattttctcaagaaggagaaataagataagaaagtcatagaa 180  
|||||  
DB 149 GTGCGAAAAGACTCGAAATTTTCTCAAGAGGAGAAATAGAAATGAAGGTCTATAGAA 208  
|||||  
QY 181 cacatcatggaagacctgtgacacaatgcagaacagcagctgagcttcgaggaattcaltc 240  
|||||  
DB 209 CACATCATGAGAGACCTGTGACACAATGACAGACAGCTGAGCTTGGAGAGATTTCATTC 268  
|||||  
QY 241 atgtctgatggaagcttaacctgtggcctcccaagagaagatgacagaggtgagcagaggtc 300  
|||||  
DB 269 ATGCTGATGGCAGAGGCTAAGCTGTGGCTCCCAAGAGAGATGCACAGAGGTGACGAGGAC 328  
|||||  
QY 301 cctggccaccacataagcagcctcgaggagagggcaccctctaa 345  
|||||  
DB 329 CCTGGCCACCACATTAAGCCAGGCTCGGGGAGGGACCCCTTAA 373  
|||||

RESULT 13  
BF511375/c 561 bp mRNA linear EST 06-DEC-2000  
LOCUS BF511375  
DEFINITION UT-H-B14-aog-e-08-0-UT.s1 NCI\_CGAP\_Sub8 Homo sapiens cDNA clone  
IMAGE:3084782 3', mRNA sequence.  
ACCESSION BF511375  
VERSION BF511375.1 GI:11594673  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 561)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/dbp/image/image.html](http://www-bio.llnl.gov/dbp/image/image.html)  
Seq primer: M13 Forward  
POLYA-yes.

## FEATURES

source

Location/Qualifiers

```

1..561
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3084782"
/lab_host="NCI-CGAP-Sub8"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; NCI-CGAP-Sub8
is a subcloned library derived from NCI-CGAP-Sub5. The
NCI-CGAP-Sub8 library had 2.5 million recombinants. A
single-stranded DNA preparation of NCI-CGAP-Sub5 was used
as a tracer in a subtractive hybridization with a driver
comprising: a pool of clones from NCI-CGAP-Sub5 (IMAGE
clone Ids 2732833-2737415, 3068040-3069191; 25% of the
driver population), a pool of clones from NCI-CGAP-Sub4
(IMAGE clone Ids 2723592-2729326; 25% of the driver
population), NCI-CGAP-Sub6 (pool AIF-AU, IMAGE Ids
2728969-2733190; 25% of the driver population), and
NCI-CGAP-Sub7 (IMAGE Ids 3069192-3072228, 3081864-3084550
; 25% of the driver population). Subtraction was
performed as previously described (Bonaldo, Lennon &
Soares (1996): Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 791-806.
TAG-LIB-NCI-CGAP_Lm19
TAG-TISSUE-lung
TAG-SEQ-GCACAC
TAG-SEQ-GCACAC
150 c 155 g 170 t

```

BASE COUNT

86 a 150 c 155 g 170 t

ORIGIN

Query Match 99.1%; Score 341.8; DB 10; Length 561;  
Best Local Similarity 99.4%; Pred. No. 3.3e-70;  
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 atgactgtcaaatgttcgacgtggaacgcaatagagaccatcatcaaccctccac 60
DB 545 ATGACTGTCAAAATGTGCGAGCTGGAACGCAACATAGAGCATCATCAACACCTTCCAC 486
QY 61 caataactctgtgaagctgggggacccagacacctgaacagggggaattcaagaagctg 120
DB 485 CAATACTCTGTGAAGCTGGGGGACCCAGACACCTGTGAACGAGGGGAATTCAAAGAGCTT 426
QY 121 gtgcgaagaagactgtgcaaatcttctcaagaagaagaataagaataaaggtcatagaa 180
DB 425 GTGCGAAAATAATGTGCAAAATTTTTCACAGAGAGAGATAGATGAAGAGTCCATAGAA 366
QY 181 cacatcatggaagacctgacacaaatgacagaacagctgagcttcgagaggtcatc 240
DB 365 CACATCATGTGAGGAGCCTGGACCAAAATGACAGACAGCTGATCTTCGAGGAGTTCATC 306
QY 241 atgctgtagcgaggttaacctggctcccaagagaagaatgacgaaggtgacgaagc 300
DB 305 ATGCTGTATGGCGAGGCTTAACCTGGGCTCCACAGAGAAATGACAGAGGCTGACGAGGGC 246
QY 301 cctggcaaccacataagccagagcctcgaggagagcaccctctaa 345
DB 245 CCGAGCCACACCAATAGCCAGGCTCGGGGAGGAGGCACCCTCTAA 201

```

RESULT 14  
BE785947 1190 bp mRNA linear EST 20-OCT-2000  
LOCUS BE785947

DEFINITION 601477976F1 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:3880747 5',  
MRNA sequence.  
ACCESSION BE785947  
VERSION BE785947.1 GI:10207145  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1190)  
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DBP/gazdar  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LAM9648 row: g column: 20  
High quality sequence stop: 521.

## FEATURES

source

Location/Qualifiers

```

1..1190
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3880747"
/lab_host="NIH-MGC_68"
/tissue_type="large cell carcinoma"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
416 a 352 c 274 g 148 t

```

BASE COUNT

416 a 352 c 274 g 148 t

ORIGIN

Query Match 98.8%; Score 341; DB 10; Length 1190;  
Best Local Similarity 100.0%; Pred. No. 6.5e-70;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 5 ctgcgaagaatgttcgacgtggaacgcaatagagaccatcatcaaccctccac 64
DB 1 CTTGCAAAATGTGCGAGCTGGAACGCAACATAGAGCATCATCAACACCTTCCACCAT 60
QY 65 actctgtgaagctgggggacccagacacctgaacagggggaattcaagaagctgtgc 124
DB 61 ACTCTGTGAAGCTGGGGGACCCAGACACCTGTGAACAGGGGAATTCAAAAGAGCTGGTGC 120
QY 125 gaaagaatctgcaaatcttctcaagaagaagaataagaatgaagaaggtcatagaa 184
DB 121 GAAAAGATCTGCAAAATTTTTCAGAGAGAGAAATGAAGAAAGCTCATAGACACA 180
QY 185 tcatggaagacctgtgacaaatgacagaacagcagctgagcttcgagaggtcatc 244
DB 181 TCATGAGAGACTGTGACCAAAATGACAGCAAGCACTGACTTTCGAGGAGTTCATTC 240
QY 245 tgatggaagagctaaacctggcctcccaagagaatgacagaaggtgacgaagccctg 304
DB 241 TGATGGCGAGGCTTAACCTGGGCTCCACAGAGAGATGACAGAGGTGACGAGGGCCCTG 300
QY 305 gccaccacataagccagagcctcgaggagagcaccctctaa 345
DB 301 GCCACACCATTAAGCCAGGCTCGGGGAGGAGGCACCCTCTAA 341

```

RESULT 15  
B518253 500 bp mRNA linear EST 29-AUG-2001  
LOCUS B518253/c  
DEFINITION 603042106T1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5182609 3',



## mRNA sequence.

## ACCESSION

B1518253

## VERSION

B1518253.1 GI:15343045

## KEYWORDS

EST.

## SOURCE

human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 500)

## AUTHORS

NIH-MGC <http://mhc.nci.nih.gov/>.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LEML at:

<http://image.llnl.gov>

Plate: L1AM11455 row: 1 column: 02

High quality sequence start: 4

High quality sequence stop: 500.

Location/Qualifiers

## FEATURES

Source

1..500

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5182609"

/clone\_lib="NIH\_MGC\_116"

/lab\_host="DH10B"

/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA

source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2

stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

023. Note: this is a NIH MGC Library."

Search completed: September 10, 2002, 06:08:02  
Job time: 6638 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: September 10, 2002, 06:42:40 ; Search time 2062.36 Seconds  
(without alignments)  
3500.676 Million cell updates/sec

Title: US-09-806-382a-2  
345  
Perfect score: 1 atgcactgcaaaatgcgca.....tcggggagggcacccctaa 345  
Sequence:

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: GenBank:  
2: gb\_ba:  
3: gb\_bt:  
4: gb\_in:  
5: gb\_om:  
6: gb\_ov:  
7: gb\_pat:  
8: gb\_ph:  
9: gb\_pl:  
10: gb\_pr:  
11: gb\_ro:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vl:  
15: em\_ba:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_mu:  
20: em\_om:  
21: em\_or:  
22: em\_ov:  
23: em\_pat:  
24: em\_ph:  
25: em\_pl:  
26: em\_ro:  
27: em\_sts:  
28: em\_un:  
29: em\_vl:  
30: em\_htg\_hum:  
31: em\_htg\_inv:  
32: em\_htg\_other:  
33: em\_htgo\_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	345	100.0	345	6	AX038706	AX038706 Sequence
2	345	100.0	462	6	AI12029	AI12029 Macrophage
3	345	100.0	462	6	HSMRP14	X06223 Human mRNA
4	345	100.0	565	6	AX332248	AX332248 Sequence
5	345	100.0	565	6	AX333056	AX333056 Sequence
6	345	100.0	565	9	HUMCFA	M26311 Human cysti
7	345	100.0	567	9	HUM2D66A09	AF086362 Homo sapi
8	345	100.0	571	6	AR015928	AF015928 Sequence
9	345	100.0	687	6	AX017310	AX017310 Sequence
10	345	100.0	688	6	AX017255	AX017255 Sequence
11	333	96.5	336	6	AI12024	AI12024 Macrophage
12	197	57.1	1754	6	AI2032	AI2032 PARTIAL Mac
13	197	57.1	4439	6	HUMMRP14A	M21064 Human migra
14	197	57.1	4440	6	I38533	I38533 Sequence 4
15	197	57.1	4804	9	AF237581	AF237581 Homo sapi
16	197	57.1	4804	9	AF237582	AF237582 Homo sapi
17	197	57.1	114364	9	AC011666	AC011666 Human Chr
18	197	57.1	164144	9	AL591704	AL591704 Human DNA
19	197	57.1	278361	2	AL606502	AL606502 Homo sapi
20	165.8	48.1	494	10	RATMRP14	LI8948 Rattus norv
21	162.6	47.1	488	10	MUSMRP14A	M83219 Mus musculu
22	162.4	47.1	480	4	AF091849	AF091849 Oryctolag
23	150	43.5	2737	6	AI12031	AI12031 PARTIAL Mac
24	98.8	28.6	273	6	AR102853	AR102853 Sequence
25	93.2	27.0	429	4	D49548	D49548 Bos taurus
26	93.2	27.0	429	6	AR083233	AR083233 Sequence
27	93.2	27.0	429	6	AR177852	AR177852 Sequence
28	93.2	27.0	429	23	E12013	E12013 cDNA encodi
29	93.2	27.0	429	23	E12020	E12020 cDNA encodi
30	92	26.7	598	5	GG126MRP	X61200 G.gallus 12
31	89.4	25.9	405	4	AF011757	AF011757 Bos tauru
32	88.8	25.7	6991	10	MM0250496	AJ250496 Mus muscu
33	84	24.3	167014	2	AC097705	AC097705 Rattus no
34	83.8	24.3	538	5	AF077613	AF077613 Salvelinu
35	81.8	23.7	439	6	AX330549	AX330549 Sequence
36	81.8	23.7	439	9	HSS100PCB	X65614 H.sapiens m
37	81.8	23.7	516	9	BC006819	BC006819 Homo sapi
38	80.4	23.3	325	6	AX284667	AX284667 Sequence
39	80.4	23.3	602	6	AX285010	AX285010 Sequence
40	79	22.9	292	6	AX341277	AX341277 Sequence
41	78.8	22.8	114	4	RABMRP14	D17404 Rabbid mRNA
42	78.6	22.8	288	6	AX038716	AX038716 Sequence
43	78.6	22.8	288	6	AX038718	AX038718 Sequence
44	78.4	22.7	289	6	AX351438	AX351438 Sequence
45	77.4	22.4	441	6	AR083238	AR083238 Sequence

## ALIGNMENTS

RESULT 1	AX038706	345 bp	DNA	linear	PAT 16-NOV-2000
LOCUS	AX038706	Sequence 17 from Patent WO0061742.			
DEFINITION	AX038706				
ACCESSION	AX038706.1	GI:11228054			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
JOURNAL	Katus/H.A. and Remppis, A.				
FEATURES	Treatment of cardiac power failure				
Source	Patent: WO 0061742-A 17 19-OCT-2000;				
	KATUS HUGO A (DE) ; REMPPIS ANDREW (DE)				
	Location/Qualifiers				
	1..345				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	1..345				
	/note="S100A9 cDNA"				
	/codon_start=1				

RESULT	2
A12029	
LOCUS	A12029 462 bp DNA linear PAT 11-JAN-1994
DEFINITION	Macrophage migration inhibition factor (MRP-14) from mRNA of a
ACCESSION	A12029
VERSION	A12029
KEYWORDS	AI2029.1 GI:490103
SOURCE	.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Kumariyoti; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 462)
TITLE	Olink,K.G., Clerc,R., Cerletti,N., Bruggen,J., Tarsay,L., Sorg,C. and Wlesendanger,W.
JOURNAL	Novel lymphokine related peptides Patent: EP 0263072-A 14 06-APR-1988;
FEATURES	CIBA-GEIGY AG
source	Location/Qualifiers
gene	1..462
CDS	/organism="Homo sapiens" 45..389 /db_xref="taxon:9606" /gene="MRP-14" 45..389 /gene="MRP-14" /codon_start=1 /protein_id="CAA01002.1" /db_xref="GI:490104" /db_xref="SWISS-Prot:P06702" /translation="MTCKASQLEKRIETIITNFHYSVKIGHPTLNGCFEKLVRKD LNKKFLKENKEVVIETHIMDDLTADKQLSEEFITLMARLTWASHKEMHEDEPGG HNHKPLKGEGTP"
BASE COUNT	139 a 128 c 124 g 71 t
ORIGIN	

FEATURES	source
LOCUS	HSMPR14
DEFINITION	Human mRNA for calcium-binding protein in macrophages (MRP-14)
ACCESSION	X06233
VERSION	Y00286
KEYWORDS	calcium binding protein.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 462)
TITLE	Odink, K.
JOURNAL	Direct Submission
REFERENCE	Submitted (12-NOV-1987) Odink K., Ciba-Geigy, Dept. of Biotechnology, K 681-442, 4002 Basel, Switzerland
AUTHORS	2 (bases 1 to 462)
TITLE	Odink, K., Cerletti, N., Bruggen, J., Clerc, R. G., Tarssey, L., Zwadlo, G., Gerhards, G., Schlegel, R. and Sory, C.
JOURNAL	Two calcium-binding proteins in infiltrate macrophages of rheumatoid arthritis
MEDLINE	Nature 330 (6143), 80-82 (1987)
COMMENT	MRP-14 is expressed myeloid specific in acute and chronic inflammation. It is increased in serum of cyclic fibrosis patients and heterozygotes for this disease see X06234.
FEATURES	Location/Qualifiers
source	1..462
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/cell_type="peripheral blood mononuclear cells"
	45..389
	/note="MRP-14 (AA 1-114)"
	/codon_start=1
	/protein_id="CAA29579.1"
	/db_xref="GI:34771"
	/db_xref="SWISS-PROT:P06702"
	/translation="MTCKMSOLERNIEITITNFQYSVKLGHPDTLNGCFKELVRKD LQNFLENNKVEVIEHIEDDTNADKQLSFEFFIMLMRLTWASHEKMEDESGP HHHKPGIGESTP"

BASE COUNT 139 a 128 c 124 g 71 t  
ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 345; DB 9; Length 462;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 atgacttcgaaatgtcgagctgtgaaagcacaatagagaccatcatcaaaccttcac 60  
|||||  
45 ATGACTTGC AAAATGTCCAGCTGGACGCAACATGAGACCATCATCAACACCTTCCAC 104  
|||||  
61 caatactctgtgaagctgggagccacagacacctgaaccaggagggaattcaaaagctg 120  
|||||  
105 CAATTA CTCTGTAGAGCTGGGACCCAGACACCCCTGAACCGAGGGGGAATTC AAAAGAGCTG 164  
|||||  
121 gtgcgaaagatctgcgaaattttctcaagaagagaatgaatgaagagtcataagaa 180  
|||||  
165 GTGCGAAAAGATCTGCAAAATTTTCTCAAGAAAGAGAAATGAATAAGGTCTCATAGAA 224  
|||||  
181 cacatcatgtgaagacctgtgacacaatgtgaagaagcgttagcttcgaagagttcatc 240  
|||||  
225 CACATCATGTGAGAGACCTGGACACAATAATGCAAGACAGCTTCCAGAGATTTCATC 284  
|||||  
241 atgctgatgtcgaggttaacctgtggcctcccaagaaatgtcacgaggttgcagagggc 300  
|||||  
285 ATGCTGATGCGGAGGCTTAACCTGGGCTCCCAAGAAAGATGCAGAGGCTGACGAGGCG 344  
|||||  
301 cctggccacacacataagccaggcctcgggaggagccacctaa 345  
|||||  
345 CCGTGGCCACCACCATTAAGCCAGGCTCTGGGAGGAGGACCCCTTAA 389  
|||||

RESULT 4  
AX332248 565 bp DNA linear PAT 09-JAN-2002  
LOCUS Sequence 2757 from Patent WO0194629.  
DEFINITION AX332248  
ACCESSION AX332248  
VERSION AX332248.1 GI:18122882  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,  
Horrihan,S., Soppet,D.R. and Weaver,Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 2757 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source 1..565  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 158 a 158 c 156 g 93 t  
ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 345; DB 6; Length 565;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 atgacttcgaaatgtcgagctgtgaaagcacaatagagaccatcatcaaaccttcac 60  
|||||  
32 ATGACTTGC AAAATGTCCAGCTGGACGCAACATAGAGACCATCAACACCTTCCAC 91  
|||||  
61 caatactctgtgaagctgggagccacagacacctgaaccaggagggaattcaaaagctg 120  
|||||  
92 CAATTA CTCTGTGAGAGCTGGGACCCAGACACCCCTGAACCGAGGGGGAATTTT AAAAGAGCTG 151  
|||||  
121 gtgcgaaagatctgcgaaattttctcaagaagagaatgaatgaagagtcataagaa 180  
|||||

Db 152 GTGCGAAAAGATCTGCAAAATTTTCTCAAGAGAGACATTAAGATCAAAAGGTCATAGAA 211  
Qy 181 cacatcatgtgaagacctgtgacacaatgtcagaagcagctgagcttcgaagagttcatc 240  
|||||  
Db 212 CACATCATGTGAGAGACCTGGACACAATAATGCAAGACAGCTGAGTTCGAGGAGTTTCATC 271  
Qy 241 atgctgatgtcgaggttaacctgtggcctcccaagaaatgtcacgaggttgcagagggc 300  
|||||  
Db 272 ATGCTGATGCGGAGGCTTAACCTGGGCTCCCAAGAAAGATGCAGAGGCTGACGAGGCG 331  
Qy 301 cctggccacacacataagccaggcctcgggaggagccacctaa 345  
|||||  
Db 332 CCGTGGCCACCACCATTAAGCCAGGCTCTGGGAGGAGGACCCCTTAA 376  
|||||

RESULT 5  
AX333056 565 bp DNA linear PAT 09-JAN-2002  
LOCUS Sequence 3565 from Patent WO0194629.  
DEFINITION AX333056  
ACCESSION AX333056  
VERSION AX333056.1 GI:18123690  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,  
Horrihan,S., Soppet,D.R. and Weaver,Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 3565 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source 1..565  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 158 a 158 c 156 g 93 t  
ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 345; DB 6; Length 565;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 atgacttcgaaatgtcgagctgtgaaagcacaatagagaccatcatcaaaccttcac 60  
|||||  
32 ATGACTTGC AAAATGTCCAGCTGGACGCAACATTAAGAGACCATCATCAACACCTTCCAC 91  
|||||  
61 caatactctgtgaagctgggagccacagacacctgaaccaggagggaattcaaaagctg 120  
|||||  
92 CAATTA CTCTGTGAGAGCTGGGACCCAGACACCCCTGAACCGAGGGGGAATTTT AAAAGAGCTG 151  
|||||  
121 gtgcgaaagatctgcgaaattttctcaagaagagaatgaatgaagagtcataagaa 180  
|||||  
152 GTGCGAAAAGATCTGCAAAATTTTCTCAAGAAAGAGAAATGAATAAGGTCTCATAGAA 211  
|||||  
181 cacatcatgtgaagacctgtgacacaatgtcagaagaagcgttagcttcgaagagttcatc 240  
|||||  
212 CACATCATGTGAGAGACCTGGACACAATAATGCAAGACAGCTTCCAGAGAGTTTCATC 271  
|||||  
241 atgctgatgtcgaggttaacctgtggcctcccaagaaatgtcacgaggttgcagagggc 300  
|||||  
272 ATGCTGATGCGGAGGCTTAACCTGGGCTCCCAAGAAAGATGCAGAGGCTGACGAGGCG 331  
|||||  
301 cctggccacacacataagccaggcctcgggaggagccacctaa 345  
|||||  
332 CCGTGGCCACCACCATTAAGCCAGGCTCTGGGAGGAGGACCCCTTAA 376  
|||||

RESULT 6  
HUMCFA HUMCFA 565 bp mRNA linear PRI 19-JUN-1995  
LOCUS

```

DEFINITION Human cystic fibrosis antigen mRNA, complete cds.
ACCESSION M26311
VERSION M26311.1 GI:862619
KEYWORDS cystic fibrosis antigen; protein kinase inhibitor.
SOURCE Homo sapiens spleen cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 565)
AUTHORS Murao,S., Collart,F.R. and Huberman,E.
TITLE A protein containing the cystic fibrosis antigen is an inhibitor of
JOURNAL protein kinases
MEDLINE J. Biol. Chem. 264 (14), 8356-8360 (1989)
COMMENT 89255276
FEATURES
    source
        location/Qualifiers
            1..565
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /tissue_type="spleen"
                32..376
                    /codon_start=1
                    /product="cystic fibrosis antigen"
                    /protein_id="AA068480.1"
                    /db_xref="GI:516621"
                    /translation="MTCKMSQLEPNIETIINHOYVYKLGHPDNLNGEPELYRKT
                    LONPLKKNKNEKVIIEHIMEDLDTNADKOLSEEFITMLARLTWASHEKMHGDEGPG
                    HHHPGIGESTP"
BASE COUNT 158 a 158 c 156 g 93 t
ORIGIN
Query Match 100.0%; Score 345; DB 9; Length 565;
Best Local Similarity 100.0%; Pred. No. 1.2e-73;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 atgacttgcacaaatgtcgcagctggaacgaacatagagaccatcatcaacaccttcac 60
DB 32 ATGACTTGCACAAATGTGCGACGTGGAGCGACACATAGACACATCATCACACCTTCCAC 91
OY 61 caatctctgtgaagctgtgggagccagacacacctgaacacagaggaatcaagaagctg 120
DB 92 CAATCTCTGTGAAGCTGTGGGACCCGACACCTGTGACACAGGGGCAATTCAGAGAGCTG 151
OY 121 gtgcgaaagaatctgcacaaatcttctcaagaagaagaataagaatgaagaagtcatalaa 180
DB 132 GTGCGAAAGAATCTGCAAAATTTCTCAGAGAGGAATGAATGAAGAAGTCATAGAA 211
OY 181 cacatcgtgagagagcctgcgacacaaatgcagacagagctgagcttcgagaggttcac 240
DB 212 CACATCATGAGAGAGCTGGACACAAATGCAACAGCAGCTGAGCTTCGAGAGTTTCATC 271
OY 241 atgctgatgctgagcttaacctggcctcccaagagaagaatgaacgaggtgacgagagc 300
DB 272 ATGCTGATGCGAGAGCTTAACCTGGGCTCCACAGAGAGATGACAGAGGCTGACGAGGCG 331
OY 301 cctggccacacccaatgaacgagagctgtgggagagagagagagagagagagagagagag 345
DB 332 CCTGGCCACACCAATGAAGCCAGGCTGTGGGAGGAGGACCCCTTAA 376

```

```

AUTHORS Moessner,J., Tan,F., Marra,M., Kucaba,T., Vandell,M., Martin,J.,
Marth,G., Bowles,L., Mylle,T., Bowers,Y., Steptoe,M., Theising,B.,
Geisels,S., Allen,M., Underwood,K., Chappell,J., Person,B.,
Gibbons,M., Harvey,N., Page,D., Chamberlain,A., Morales,R.,
Schurk,R., Ritzer,E., Kohn,S., Swaller,T., Behymer,K., Hillier,L.,
Wilson,R. and Waterston,R.
TITLE Full Clone Sequencing of the Longest Available Member from Each
JOURNAL Unigene Cluster
REFERENCE 2 (bases 1 to 567)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
SUBMITTED BY:
Genome Sequencing Center
Department of Genetics
Washington University
St. Louis MO 63108, USA
http://genome.wustl.edu/gsc
mailto:est@watson.wustl.edu
NOTICE: This sequence represents the full insert of this cDNA. No
attempt has been made to verify whether this corresponds to the
full-length of the original mRNA from which it was derived. We
have tried to obtain double-stranded, or double chemistry sequence
across the entire clone, but potentially, there are areas in the
sequence where this level of coverage was not achieved.
Nevertheless, we are confident of the accuracy of this sequence as
all regions of low quality, as defined by PHRAP (P. Green, in
preparation), were visually inspected and edited accordingly. The
consensus quality values for this sequence have been submitted
separately.
SIMILARITY INFORMATION:
Probable match to Homo sapiens protein P06702 (PID:g115444)
CALGRANDULIN B (MIGRATION INHIBITORY FACTOR-RELATED PROTEIN 14)
(MRP-14) (P14) (LEUKOCYTE LI COMPLEX HEAVY CHAIN) (S100
CALCIUM-BINDING PROTEIN A9)
FEATURES
    source
        location/Qualifiers
            1..567
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:345592"
                /clone_1ib="Soares_fetal_heart_NbHH19W"
                25..366
                    /note="probable match to Homo sapiens protein P06702
                    (PID:g115444)"
                37..384
                    /note="similar to Bos taurus protein A42628
                    (PID:g1072439)"
                37..384
                    /note="similar to Bos taurus protein A42628
                    (PID:g1072439)"
                37..384
                    /note="similar to Bos taurus protein P28783 (PID:g399172)"
                40..357
                    /note="similar to Mus musculus protein P31725
                    (PID:g399173)"
                40..360
                    /note="similar to Rattus norvegicus protein P50116
                    (PID:g1710814)"
BASE COUNT 166 a 157 c 154 g 90 t
ORIGIN
Query Match 100.0%; Score 345; DB 9; Length 567;
Best Local Similarity 100.0%; Pred. No. 1.2e-73;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 atgacttgcacaaatgtcgcagctggaacgaacatagagaccatcatcaacaccttcac 60
DB 25 ATGACTTGCACAAATGTGCGACGTGGAGCGACACATAGACACATCATCACACCTTCCAC 84

```

QY 61 caatactctgtgaagctggggcaccagacacccctgaagcagggggaattcaagagctg 120  
 Db 85 CAATACCTCTGTGAAGCTGGGGCACCCAGACACCTGAAACCGGGGGAATTCAAAGAGCTG 144  
 QY 121 gtgcgaaagaatctgcgaatttctcaagaagaagaataagaatgaagaagtcataagaa 180  
 Db 145 GTGGGAAAGATCTGCAAAATTTTCTCAAGAGAGAGAAATGAATGAAGAGTCATAGAA 204  
 QY 181 cacatcatgagagaccttgcagacaatgacagaagcagctgaagcttcgagaggttcac 240  
 Db 205 CACATCATGAGAGACCTGAGACAAATGACAGACAAGCAGCTGACCTTCGAGAGATTCATC 264  
 QY 241 atgtctgtagggagagctgaagctgggtcccaagagaagaatgcaagaggtgaagagcc 300  
 Db 265 ATGCTGATGGGAGAGCTAACTGGGCGCTCCACAGAGAAGATGACAGAGGTGACGAGGCG 324  
 QY 301 cctggccaccacataagccagagcctggggagagggaccacctaa 345  
 Db 325 CCTGGCCACACCATTAAGCCAGGCGCTCGGGAGGAGGACCCCTTAA 369

RESULT 8  
 ARO15928  
 LOCUS ARO15928 571 bp DNA linear PAT 05-DEC-1998  
 DEFINITION Sequence 4 from patent US 5776348.  
 ACCESSION ARO15928  
 VERSION ARO15928.1 GI:3972205  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 571)  
 AUTHORS Selengut,J.D., Orme-Johnson,W.H., Dretler,S.P. and Asakura,H.  
 TITLE Mineral precipitation system and method for inhibiting mineral  
 JOURNAL precipitate formation  
 FEATURES Patent: US 5776348-A 4 07-JUL-1998;  
 LOCATION/Qualifiers  
 SOURCE 1. .571  
 BASE COUNT 160 a 160 c 157 g 94 t  
 ORIGIN

Query Match 100.0%; Score 345; DB 6; Length 571;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-73;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgacttgcgaatgtgcagctggaagcaacataagagaccatcatcaaccttcac 60  
 Db 44 ATGACTTGCAGAAATGTGCGAGCTGGAACGCAACATGAGACCATCATCAACCTTCCAC 103  
 QY 61 caatactctgtaagctgggacacacacacccctgaacccagggggaattcaagaagctg 120  
 Db 104 CAATACCTCTGTGAAGCTGGGGCACCCAGACACCTGAAACCGGGGAATTAACAGCTG 163  
 QY 121 gtgcgaaagaatctgcgaatttctcaagaagaagaataagaatgaagaagtcataagaa 180  
 Db 164 GTGGGAAAGATCTGCAAAATTTTCTCAAGAGAGAGAAATGAATGAAGAGTCATAGAA 223  
 QY 181 cacatcatgagagaccttgcagacaatgacagaagcagctgaagcttcgagaggttcac 240  
 Db 224 CACATCATGAGAGACCTGAGACAAATGACAGACAAGCAGCTGAGCTTCGAGAGATTCATC 283  
 QY 241 atgtctgtagggagagctgaagctgggtcccaagagaagaatgcaagaggtgaagagcc 300  
 Db 284 ATGCTGATGGGAGAGCTAACTGGGCGCTCCACAGAGAAGATGACAGAGGTGACGAGGCG 343  
 QY 301 cctggccaccacataagccagagcctggggagagggaccacctaa 345  
 Db 344 CCTGGCCACACCATTAAGCCAGGCGCTCGGGAGGAGGACCCCTTAA 388

RESULT 9

AX017310  
 LOCUS AX017310 687 bp DNA linear PAT 07-SEP-2000  
 DEFINITION Sequence 66 from Patent WO947669.  
 ACCESSION AX017310  
 VERSION AX017310.1 GI:10042228  
 KEYWORDS human.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 687)  
 AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and  
 Pilarsky,C.  
 TITLE Human nucleic acid sequences from tissue of breast tumors  
 JOURNAL Patent: WO 947669-A 66 23-SEP-1999;  
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUNG  
 (DE); PILARSKY CHRISTIAN (DE)  
 FEATURES Location/Qualifiers  
 SOURCE 1. .687  
 BASE COUNT 177 a 198 c 190 g 122 t  
 ORIGIN

Query Match 100.0%; Score 345; DB 6; Length 687;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-73;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgacttgcgaatgtgcagctggaagcaacataagagaccatcatcaaccttcac 60  
 Db 101 ATGACTTGCAGAAATGTGCGAGCTGGAACGCAACATGAGACCATCATCAACCTTCCAC 160  
 QY 61 caatactctgtaagctgggacacacacacccctgaacccagggggaattcaagaagctg 120  
 Db 161 CAATACCTCTGTGAAGCTGGGGCACCCAGACACCTTAAACCGGGGAATTCAAAGAGCTG 220  
 QY 121 gtgcgaaagaatctgcgaatttctcaagaagaagaataagaatgaagaagtcataagaa 180  
 Db 221 GTGGGAAAGATCTGCAAAATTTTCTCAAGAGAGAGAAATGAATGAAGAGTCATAGAA 280  
 QY 181 cacatcatgagagaccttgcagacaatgacagaagcagctgaagcttcgagaggttcac 240  
 Db 281 CACATCATGAGAGACCTGAGACAAATGACAGAACAGCTGAGCTTCGAGAGATTCATC 340  
 QY 241 atgtctgtagggagagctgaagctgggtcccaagagaagaatgcaagaggtgaagagcc 300  
 Db 341 ATGCTGATGGGAGAGCTAACTGGGCGCTCCACAGAGAAGATGACAGAGGTGACGAGGCG 400  
 QY 301 cctggccaccacataagccagagcctcggggagggaccacctaa 345  
 Db 401 CCTGGCCACACCATTAAGCCAGGCGCTCGGGAGGAGGACCCCTTAA 445

RESULT 10  
 AX017255  
 LOCUS AX017255 688 bp DNA linear PAT 07-SEP-2000  
 DEFINITION Sequence 5 from Patent WO947669.  
 ACCESSION AX017255  
 VERSION AX017255.1 GI:10042173  
 KEYWORDS human.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 688)  
 AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and  
 Pilarsky,C.  
 TITLE Human nucleic acid sequences from tissue of breast tumors  
 JOURNAL Patent: WO 947669-A 5 23-SEP-1999;  
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUNG

(DE); PILARSKY CHRISTIAN (DE)  
FEATURES  
source 1..688  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 178 a 198 c 190 g 122 t  
ORIGIN

Query Match 100.0%; Score 345; DB 6; Length 688;  
Best Local Similarity 100.0%; Pred. No. 1.2e-73;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgacctcaaatgctgcagctggaacgacaatagagacatcatcaaaccttccac 60  
|||  
Db 101 ATGACTTCAAAATGCTGCAGCTGGAACGCAACATAGAGACCATCATCAACCTTCCAC 160

QY 61 caatctctgtgaagctgaggacaccagacacccctgaaccaggggaaatlcaaaagctg 120  
|||  
Db 161 CAATACTCTGTGAAGCTGGGACACCAGACACCTGAACAGGGGAATTCAAAGAGCTG 220

QY 121 gtggaaaaagatctgcgaatttctcagaagagagaataagaataaggtctatgaa 180  
|||  
Db 221 GTGGAAAAAGATCTGCAAAATTTTCTCAAGAAAGGAATGAAGTGAAGGTCATTAGAA 280

QY 181 cacatcatgagagacctgagacacaatgacagaacagcagctgagcttgcagagatcatc 240  
|||  
Db 281 CACATCATGAGAGAGACTGAGACACAATGCAGACAGAGCTGTGAGAGAGTTCAATC 340

QY 241 atgctgtagggagagctaaacctgggctcccaagagaagatgcagaggtgaacgagggc 300  
|||  
Db 341 ATGCTGAGGGGAGGAGCTTAACCTGGGCCCTCCACGAGAAAGATGACAGAGGTGACGAGGGC 400

QY 301 cctggccaccaccataagccagagcctcgaggaggaccccttaa 345  
|||  
Db 401 CCTGGCCACACCATTAAGCCAGGCTCGGGAGGAGGCCCTCTAA 445

RESULT 11  
A12024 336 bp DNA linear PAT 11-JAN-1994  
LOCUS  
DEFINITION Macrophage migration inhibition factor (MRP-14).  
ACCESSION A12024  
VERSION A12024.1 GI:491245  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 336)  
AUTHORS Odink,K.G., Clerc,R., Cerletti,N., Brueggen,J., Tarcsay,L., Sorg,C.  
and Wiesendanger,W.  
TITLE Novel lymphokine related peptides  
JOURNAL Patent: EP 0263072-A 9 06-APR-1988;  
CIBA-GEIGY AG

FEATURES  
source 1..336  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
misc\_difference 1 /note="n= a flanking DNA residue of 12 nucleotides or more  
containing a promoter sequence"  
misc\_difference 2 /note="n= M,T,C,K or is absent"  
gene 3..335  
/gene="MRP-14"  
3..335  
/gene="MRP-14"  
/codon\_start=1  
/protein\_id="CAA01004.1"  
/transl\_table=1  
/protein\_id="CAAO0399.1"  
/db\_xref="GI:491246"  
/translation="MSOLERNITITINTFHQYSVKLGHPDTLNOGEEKELVRDLQNF  
LKRNKNEKVIETIHMEDLDTNADKQLSFEETIMLARLTWASHKHEKHEGDEGPGHHHK

PGIGGCTP"  
misc\_difference 336  
/note="n= a flanking DNA residue of one or more  
nucleotides or is absent"  
BASE COUNT 105 a 87 c 90 g 51 t 3 others  
ORIGIN

Query Match 96.5%; Score 333; DB 6; Length 336;  
Best Local Similarity 100.0%; Pred. No. 1e-70;  
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 atgtgcagctggaacggaacatagagaccatcatcaaaccttcacaaactctgtg 72  
|||  
Db 3 ATGTGCGAGCTGGAACGCAACATAGAGACCATCATCAACACCTTCACCAATACTCTGTG 62

QY 73 aagctggggcaccacagaaacccctgaaccagggggaattcaaaagctgtgtcgaaaagat 132  
|||  
Db 63 AAGCTGGGGCACACCAGACACCTGAACAGGGGGAATTCAAAAGAGCTGTGCAAAAGAT 122

QY 133 ctgcgaattttctcaaaagagagataagaatgaaaggtcatatgacacatcatgag 192  
|||  
Db 123 CTGCAAAATTTTCTCAAGAAGAGATTAAGATGAAGTGAAGTCAATGAAACACATCATGGAG 182

QY 193 gacctggacacaatgacgaacaagcagctgagcttcgagaggtcatcatgtgtatggcg 252  
|||  
Db 183 GACCTGGACACAAATGACAGACAGCAGCTGAGCTTCAGAGAGTTTCATCTGTATGGCG 242

QY 253 aggttaacctgggctcccaegagaagatgcagaggtgtacgagggccttggccaacac 312  
|||  
Db 243 AGGCTTAACCTGGGCTCCACAGAGAAGATGCGAGGGGTGACGAGGGCTGGCCACAC 302

QY 313 cataagcagagcctcgaggaggagaccccttaa 345  
|||  
Db 303 CATTAAGCCAGGCTCGGGAGGAGGCCACCCCTTA 335

RESULT 12  
A12032 1754 bp DNA linear PAT 12-JAN-1994  
LOCUS  
DEFINITION PARTIAL Macrophage migration inhibition factor (MRP-14) from a  
human Placenta or fetal liver cells (formula 6).  
ACCESSION A12032  
VERSION A12032.1 GI:490107  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1754)  
AUTHORS Odink,K.G., Clerc,R., Cerletti,N., Brueggen,J., Tarcsay,L., Sorg,C.  
and Wiesendanger,W.  
TITLE Novel lymphokine related peptides  
JOURNAL Patent: EP 0263072-A 18 06-APR-1988;  
CIBA-GEIGY AG

FEATURES  
source 1..1754  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
gene 711..905  
/gene="MRP-14"  
711..905  
/partial  
/gene="MRP-14"  
/codon\_start=1  
/protein\_id="CAA01004.1"  
/db\_xref="GI:490108"  
/db\_xref="SWISS-PROT:P06702"  
/translation="KENKNEKVIETIHMEDLDTNADKQLSFEETIMLARLTWASHKHEKHEGDEGSGHHKRGIGGCTP"



Query Match  
Best Local Similarity 57.1%; Score 197; DB 6; Length 1754;  
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 agaaagagaataaagatgaagatcatagacacatcatggaagacctggaacacaatg 208  
|||||  
Db 709 AGAAGGAGAAATAGAAATGAAAAGCTCATAGAACATCATCATGAGAGCTGAGACACAAATG 768

QY 209 cagacaagacagctgagctcgaagagatcatcatgctgagtcgaagcctaaccctggagcct 268  
|||||  
Db 769 CAGACAAACGACTGAGCTTCGAGGAGTTCATCATGCTGATGGCAGGCTAACCTGGGCT 828

QY 269 cccacgagaagatgcacagaggtgacagagggccctggccaccacataagcagagcctg 328  
|||||  
Db 829 CCCACGAGAAAGATGCACAGAGGTGACGAGGGCCTGGCCACCATTAAGCCAGGCTCG 888

QY 329 gggagggcaccacctaa 345  
|||||  
Db 889 GGGAGGGCACCCCTTAA 905

RESULT 13  
LOCUS HUMMRP14 4439 bp DNA linear PRI 27-APR-1993  
DEFINITION Human migration inhibitory factor-related protein 14 (MRP14) gene,  
complete cds.  
ACCESSION M21064  
VERSION M21064.1 GI:188689  
KEYWORDS migration inhibitory factor-related protein.  
SOURCE Human blood monocyte cell, clone pUCMRP14.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 4439)  
AUTHORS Lagasse, E. and Clerc, R.G.  
TITLE Cloning and expression of two human genes encoding calcium-binding  
proteins that are regulated during myeloid differentiation  
JOURNAL Mol. Cell. Biol. 8, 2402-2410 (1988)  
MEDLINE 88302148  
FEATURES  
source Location/Qualifiers  
1..4439  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
prim\_transcript 1001..3825  
/note="MRP14 mRNA and introns"  
intron 1029..1415  
/note="MRP14, intron A"  
join(1431..1580,3448..3642)  
/partial  
/note="migration inhibitory factor-related protein 14"  
/codon\_start=1  
/protein\_id="AA36326.1"  
/db\_xref="GI:386958"  
/translation="MTCKMSOLERNITETIIFPHOYSVKLGHPDPTLNOGEPKELVYKD  
LONELKKNENKIVIEIHMEDLDPTNADKLSFEFTIMLARLTWASHHEKMHGDESGP  
HHKRPGLGEGTP"  
exon <1431..1580  
/number=2  
/note="migration inhibitory factor-related protein 14"  
intron 1581..3447  
/note="MRP14, intron B"  
exon 3448..3642  
/number=3  
/note="migration inhibitory factor-related protein 14"  
BASE COUNT 1099 a 1203 c 1159 g 978 t  
ORIGIN

Query Match  
Best Local Similarity 57.1%; Score 197; DB 9; Length 4439;  
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 agaaagagaataaagatgaagatcatagacacatcatggaagacctggaacacaatg 208  
|||||  
Db 3446 AGAAGGAGAAATAGAAATGAAAAGCTCATAGAACATCATCATGAGAGCTGAGACACAAATG 3505

QY 209 cagacaagacagctgagctcgaagagatcatcatgctgagtcgaagcctaaccctggagcct 268  
|||||  
Db 3506 CAGACAAACGACTGAGCTTCGAGGAGTTCATCATGCTGATGGCAGGCTAACCTGGGCT 3565

QY 269 cccacgagaagatgcacagaggtgacagagggccctggccaccacataagcagagcctg 328  
|||||  
Db 3566 CCCACGAGAAAGATGCACAGAGGTGACGAGGGCCTGGCCACCATTAAGCCAGGCTCG 3625

QY 329 gggagggcaccacctaa 345  
|||||  
Db 3626 GGGAGGGCACCCCTTAA 3642

RESULT 14  
LOCUS I38533 4440 bp DNA linear PAT 13-MAY-1997  
DEFINITION Sequence 4 from patent US 5614397.  
ACCESSION I38533  
VERSION I38533.1 GI:2084587  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4440)  
AUTHORS Weissman, I. and Lagasse, E.  
TITLE Method and compositions for modulating lifespan of hematology  
cells  
JOURNAL Patent: US 5614397-A 4 25-MAR-1997;  
FEATURES  
source Location/Qualifiers  
1..4440  
/organism="unknown"  
BASE COUNT 1099 a 1203 c 1159 g 978 t 1 others  
ORIGIN

Query Match  
Best Local Similarity 57.1%; Score 197; DB 6; Length 4440;  
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 agaaagagaataaagatgaagatcatagacacatcatggaagacctggaacacaatg 208  
|||||  
Db 3447 AGAAGGAGAAATAGAAATGAAAAGCTCATAGAACATCATCATGAGAGCTGAGACACAAATG 3506

QY 209 cagacaagacagctgagctcgaagagatcatcatgctgagtcgaagcctaaccctggagcct 268  
|||||  
Db 3507 CAGACAAACGACTGAGCTTCGAGGAGTTCATCATGCTGATGGCAGGCTAACCTGGGCT 3566

QY 269 cccacgagaagatgcacagaggtgacagagggccctggccaccacataagcagagcctg 328  
|||||  
Db 3567 CCCACGAGAAAGATGCACAGAGGTGACGAGGGCCTGGCCACCATTAAGCCAGGCTCG 3626

QY 329 gggagggcaccacctaa 345  
|||||  
Db 3627 GGGAGGGCACCCCTTAA 3643

RESULT 15  
LOCUS AF237581 4804 bp DNA linear PRI 05-APR-2000  
DEFINITION Homo sapiens migration inhibitory factor-related protein 14 variant  
P (S100A9) gene, complete cds.  
ACCESSION AF237581  
VERSION AF237581.1 GI:7417326  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 4804)

**AUTHORS** Wang, M.-R., Xu, X., Cai, Y., Xu, H., Han, Y.-L., Xu, Z.-X. and Wu, M.  
**TITLE** Human gene for migration inhibitory factor-related protein 14 (MRP14), variant allele  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 4804)  
**AUTHORS** Wang, M.-R., Xu, X., Cai, Y., Xu, H., Han, Y.-L., Xu, Z.-X. and Wu, M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (20-FEB-2000) National Laboratory of Molecular Oncology, Cancer Institute, CAMS, PUMC, Panjiayuan, Chaoyang Qu, Beijing 100021, China

**FEATURES**  
**source** Location/Qualifiers  
 1..4804  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 join(1419..1584,3796..4176)  
 /gene="S100A9"  
 /product="migration inhibitory factor-related protein 14 variant p"  
 1419..4176  
 /gene="S100A9"  
 /note="MRP14"  
 join(1435..1584,3796..3990)  
 /gene="S100A9"  
 /note="MRP14p"  
 /codon\_start=1  
 /product="migration inhibitory factor-related protein 14 variant p"  
 /protein\_id="AA62536.1"  
 /db\_xref="GI:7417327"  
 /translation="MTCKMSQLERNIETIINTFHQYSVKLGHPDTLNGEKKELVRKD LQNFLEKKNKMEKYIEHIMEDLDNADKQLSFEFIMLARLTWASHEKMHGDEGPG HHHRPGLEGTP"

**CDS**  
 join(1435..1584,3796..3990)  
 /gene="S100A9"  
 /note="MRP14"  
 /codon\_start=1  
 /product="migration inhibitory factor-related protein 14 variant p"  
 /protein\_id="AA62536.1"  
 /db\_xref="GI:7417327"  
 /translation="MTCKMSQLERNIETIINTFHQYSVKLGHPDTLNGEKKELVRKD LQNFLEKKNKMEKYIEHIMEDLDNADKQLSFEFIMLARLTWASHEKMHGDEGPG HHHRPGLEGTP"

**gene**  
 join(1435..1584,3796..3990)  
 /gene="S100A9"  
 /note="MRP14"  
 /codon\_start=1  
 /product="migration inhibitory factor-related protein 14 variant p"  
 /protein\_id="AA62536.1"  
 /db\_xref="GI:7417327"  
 /translation="MTCKMSQLERNIETIINTFHQYSVKLGHPDTLNGEKKELVRKD LQNFLEKKNKMEKYIEHIMEDLDNADKQLSFEFIMLARLTWASHEKMHGDEGPG HHHRPGLEGTP"

**BASE COUNT** 1158 a 1317 c 1237 g 1092 t

**ORIGIN**

Query Match 57.1%; Score 197; DB 9; Length 4804;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-37;  
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 149 agaagagaataagaatgaagaagtcataagaacacatcatgtgagagaccttgacacaaatg 208  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
**Db** 3794 AGAAGAGAAATGAAGATGAAGATGATGAACACATCATGTGAGAGACCTGACACAAATG 3853

**QY** 209 cagacaagcagctgagcttcgagaggttcacatcatgtgagagaccttgagcct 268  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
**Db** 3854 CAGACAGCAGCTGAGCTTCGAGAGTTCATCATGTGATGGGAGGCTAACCTGGGCT 3913

**QY** 269 ccacagagaagatgacgagaggtgacgagggccttgccaccacataagccaagcctcg 328  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
**Db** 3914 CCCACGAGAAGATGCACGAGGGTGAAGGGGCGCTGGCCACCACCATAGCCAGGCGCTCG 3973

**QY** 329 gggagggcaccacctaa 345  
 ||||||||||||||||||  
**Db** 3974 GGGAGGGCACCCCTAA 3990

Search completed: September 10, 2002, 06:42:47  
 Job time: 6403 sec

OM of: US-09-806-382a-1 to: Issued\_Patents\_AA:\* out\_format : pfs  
Date: Sep 9, 2002 3:11 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL-frame+ntp.model -DEV=xlp  
-O/cgn2\_1/USPRO.spool/US09806382/runatc\_09092002\_143848\_15391/app-query.fasta\_1.752  
-DB-Issued\_Patents\_AA -OPMT=fastan -SUFFIX=ra1 -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -ICOPEXT=0.000  
-GAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bloms62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09806382 @cgn1\_139 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-09-806-382a-1  
Query length: 282  
Database: Issued\_Patents\_AA:\*  
Database sequences: 231628  
Database length: 24425594  
Search time (sec): 38.230000

## Score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation	..		
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-7			485.00	1142.41	4.1e-56	93	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-16			485.00	1142.41	4.1e-56	93	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-385-241-1			485.00	1142.41	4.1e-56	93	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-14			295.00	690.88	6.1e-31	89	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-17			247.00	668.23	1.1e-30	88	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-200-016-3			245.00	578.51	2.1e-24	46	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-200-016-2			245.00	566.84	9.5e-24	47	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-056-200-101			236.00	557.55	3.2e-23	45	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-800-644-101			236.00	557.55	3.2e-23	45	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-056-200-108			213.00	503.80	3.5e-20	41	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-800-644-108			213.00	503.80	3.5e-20	41	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-568-310D-19			160.00	369.38	4.7e-13	92	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-970-455-19			160.00	369.38	4.7e-13	92	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-568-310D-20			158.00	364.63	8.7e-13	92	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-09-270-455-20			158.00	364.63	8.7e-13	92	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-794-000-2			151.00	348.09	7.4e-12	91	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-918-727-5			139.00	319.43	2.9e-10	92	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-09-205-680A-5			139.00	319.43	2.9e-10	92	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-09-051-589-1			138.00	317.05	3.9e-10	92	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-11			137.00	314.78	5.3e-10	91	i	Sg
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-385-241-3			133.50	304.13	1.6e-09	114	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-8			131.50	299.84	3.0e-09	109	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-918-727-7			130.50	297.09	4.1e-09	113	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-18			127.00	299.99	8.4e-09	38	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-190-560-2			126.00	287.54	1.6e-08	101	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-469-946-2			126.00	287.54	1.6e-08	101	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-468-942-2			126.00	287.54	1.6e-08	101	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-12			118.50	272.77	1.4e-07	97	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-11			118.50	272.77	1.4e-07	97	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-918-727-6			111.00	251.46	1.5e-06	105	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-09-205-680A-6			111.00	251.46	1.5e-06	105	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-469-946-58			109.50	248.29	2.4e-06	101	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-469-946-58			109.50	248.29	2.4e-06	101	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-10			108.00	246.02	3.4e-06	89	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-08-918-727-1			105.50	238.57	8.2e-06	103	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-09-205-680A-1			105.50	238.57	8.2e-06	103	i	S
/cgn2_6/p/ptodata/2/1aa/5A_COMB.pep:US-07-987-272a-9			103.50	234.65	1.5e-05	95	i	S

/cgn2\_6/p/ptodata/2/1aa/5B\_COMB.pep:US-08-918-727-3 + 101.50 229.57 2.7e-05 98 i  
/cgn2\_6/p/ptodata/2/1aa/5A\_COMB.pep:US-09-205-680A-3 + 101.50 229.57 2.7e-05 98 i  
/cgn2\_6/p/ptodata/2/1aa/5A\_COMB.pep:US-09-048-889-11 + 101.50 229.57 2.7e-05 98 i  
/cgn2\_6/p/ptodata/2/1aa/5A\_COMB.pep:US-08-056-200-94 + 98.00 190.69 0.0002 1898  
/cgn2\_6/p/ptodata/2/1aa/5B\_COMB.pep:US-08-800-644-94 + 98.00 190.69 0.0002 1898

## seq\_name: /cgn2\_6/p/ptodata/2/1aa/5A\_COMB.pep:US-07-987-272a-7

## seq\_documentation\_block:

Sequence 7, Application US/07987272A

Patent No. 5731166

GENERAL INFORMATION:

APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M

TITLE OF INVENTION: No. 5731166el Chemotactic Factor

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESS: Cushman Darby & Cushman

STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower

CITY: Washington

STATE: D. C.

COUNTRY: USA

ZIP: 20005-3918

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/987, 272A

FILING DATE: 05-MAR-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PK 2127

FILING DATE: 05-FEB-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PK 4463

FILING DATE: 05-SEP-1991

ATTORNEY/AGENT INFORMATION:

NAME: Brinkman, David W

REGISTRATION NUMBER: 20,817

REFERENCE/DOCKET NUMBER: DMB/1925/200259

TELEPHONE: 202-822 3000

TELEFAX: 202-822 0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 93 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-987-272a-7

alignment\_scores:

Quality: 485.00 Length: 93

Ratio: 5.215 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-806-382a-1 x US-07-987-272a-7 ..

Align seg 1/1 to: US-07-987-272a-7 from: 1 to: 93

1 ATGTGACCGAGCGGGAAGCCCTGAACCTTCATCCAGCGCTACCA 50  
|||||  
1 MetcunrignluteGlltlyAlaleuansserllellelspvallyrhl 17  
|||||  
51 CAAGTACTCCCTGATTAAGGGGAATTTCATGCCGTACAGGAGTACC 100  
|||||  
17 slsytserleuileuileysglYanphenHlsAlaVallyrArgspaspl 34

101 TGAAGAAATTCGTAGAGACCGAGTCTCTCAGTATATCAGAGAAAAGGT 150  
|||||  
34 euylsYslsleuLeuGlutHnGlucysProglntYrIleaArgLysGly 50  
|||||  
151 GCAGAGCTGTGTTCAAGAGTTGGATATCACACTGATGTCAGTTAA 200  
|||||  
51 AlaAspValItrPheLysGluLeuAspIleAsnThrAspGlyAlaValas 67  
|||||  
201 CTTCAGAGATTCTCATTTCTGTGATTAAGATGGCGGTGCGAGCCACA 250  
|||||  
67 nPhgIngluPheLeuIleLeuValIleLysMetGlyValAlaAlaHisL 84  
|||||  
251 AAAAAAGCCATGAAGAAAGCCACCAAGAG 279  
|||||  
84 yslsSerHisGluGluSerHisLysGlu 93  
seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-987-272A-16  
seq\_documentation\_block:  
: Sequence 16, Application US/07987272A  
: Patent No. 5731166  
: GENERAL INFORMATION:  
: APPLICANT: Gezey, C., Simpson, R. J. and Lackmann, M  
: TITLE OF INVENTION: No. 5731166el Chemolactic Factor  
: NUMBER OF SEQUENCES: 23  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: Cushman Darby & Cushman  
: STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower  
: CITY: Washington  
: STATE: D. C.  
: COUNTRY: USA  
: ZIP: 20005-3918  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/07/987,272A  
: FILING DATE: 05-MAR-1993  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: AU PK 2127  
: FILING DATE: 05-FEB-1990  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: AU PK 4463  
: FILING DATE: 05-SEP-1991  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Brinkman, David W  
: REGISTRATION NUMBER: 20,817  
: REFERENCE/DOCKET NUMBER: DMB/1925/200259  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 202-861 3000  
: TELEFAX: 202-822 0944  
: TELETYPE: 6714627 CUSH  
: INFORMATION FOR SEQ ID NO: 16:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 93 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-07-987-272A-16  
alignment\_scores:  
Quality: 485.00 Length: 93  
Ratio: 5.215 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-806-382A-1 x US-07-987-272A-16 ..  
Align seg 1/1 to: US-07-987-272A-16 from: 1 to: 93

1 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATCGACGCTACCA 50  
|||||  
1 MetLeuThrGluLeuGluLysAlaLeuAsnSerIleIleAspValIlyrH 17  
|||||  
51 CAAGTACTCCCTGATTAAGGGAATTTCCATGCCGCTCTACAGGAGATACC 100  
|||||  
17 slYsYrSerLeuIleLysGlyAsnPhenHisAlaValIlyrArgAspAsp 34  
|||||  
101 TGAAGAAATTCGTAGAGACCGAGTCTCTCAGTATATCAGAGAAAAGGT 150  
|||||  
34 euylsYslsleuLeuGlutHnGlucysProglntYrIleaArgLysGly 50  
|||||  
151 GCAGAGCTGTGTTCAAGAGTTGGATATCACACTGATGTCAGTTAA 200  
|||||  
51 AlaAspValItrPheLysGluLeuAspIleAsnThrAspGlyAlaValas 67  
|||||  
201 CTTCAGAGATTCTCATTTCTGTGATTAAGATGGCGGTGCGAGCCACA 250  
|||||  
67 nPhgIngluPheLeuIleLeuValIleLysMetGlyValAlaAlaHisL 84  
|||||  
251 AAAAAAGCCATGAAGAAAGCCACCAAGAG 279  
|||||  
84 yslsSerHisGluGluSerHisLysGlu 93  
seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-385-241-1  
seq\_documentation\_block:  
: Sequence 1, Application US/08385241  
: Patent No. 5776348  
: GENERAL INFORMATION:  
: APPLICANT: Selengut Ph.D., Jeremy D.  
: APPLICANT: Orme-Johnson Ph.D., William H.  
: APPLICANT: Dretler M.D., Stephen P.  
: APPLICANT: Asakura M.D., Hirokaka  
: TITLE OF INVENTION: SYSTEM AND METHOD FOR INHIBITING  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: Choate, Hall & Stewart  
: STREET: 53 State Street  
: CITY: Boston  
: STATE: MA  
: COUNTRY: USA  
: ZIP: 02109-2891  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/385,241  
: FILING DATE:  
: CLASSIFICATION: 514  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Herschbach Ph.D., Brenda M.  
: REGISTRATION NUMBER: P-39,223  
: REFERENCE/DOCKET NUMBER: 492611-000 (MIT6915)  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (617) 248-5175  
: TELEFAX: (617) 248-4000  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 93 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: IMMEDIATE SOURCE:  
: CLONE: hmrp-8 protein  
: US-08-385-241-1

## alignment\_scores:

Quality: 485.00 Length: 93  
Ratio: 5.215 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-806-382A-1 x US-08-385-241-1 ..

Align seg 1/1 to: US-08-385-241-1 from: 1 to: 93

```
1  AAGTTGACCGAGCTGAGAAAGCCTTGACTATCATGACGCTTACCA 50
|||||
17 MetLeuThrGluLeuGlnLysAlaLeuAsnSerIleIleLeuSpValTyrH1 17
51 CAAGTACTCTCGATTAAGGGAATTTCCATGCTCTACAGGAGTAC 100
|||||
17 sLysTyrSerLeuIleLysGlyAsnPheHisAlaValTyrAspAspL 34
101 TGAAGAAATTCCTAGACCGAGTGTCTCAGTATATCAGAAAAAGGT 150
|||||
34 euLysLysLeuLeuGlnThrGlnCysProGlnTyrIleArgLysLysGly 50
151 GCAGACGCTGTGTTCAAGAGTTGGATATCAACACTGATGTGAGTTAA 200
|||||
51 AlaAspValTyrPheLysGlnLeuAspIleAsnThrAspGlyAlaValAs 67
201 CTTCCAGAGTTCCTCATCTGTTGATTAAGATGGGCGTGCAGCCCA 250
|||||
67 nPheGlnGlnPheLeuIleLeuValIleLysMetGlyAlaAlaHisL 84
251 AAAAAAGCCATGAAGAACCCACAAAGAG 279
|||||
84 ySLysSerHisGlnGlnSerHisLysGln 93
```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-987-272A-14

## seq\_documentation\_block:

; Sequence 14, Application US/07987272A  
; Patent No. 5731166  
; GENERAL INFORMATION:  
; APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M  
; TITLE OF INVENTION: No. 5731166el Chemotactic Factor  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cushman Darby & Cushman  
; STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/987,272A  
; FILING DATE: 05-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PK 2127  
; FILING DATE: 05-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PK 4463  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brinkman, David W  
; REGISTRATION NUMBER: 20,817  
; REFERENCE/DOCKET NUMBER: DMB/1925/200259  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861 3000

; TELEFAX: 202-822 0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 89 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-987-272A-14

## alignment\_scores:

Quality: 295.00 Length: 89  
Ratio: 3.734 Gaps: 0  
Percent Similarity: 88.764 Percent Identity: 58.427

## alignment\_block:

US-09-806-382A-1 x US-07-987-272A-14 ..

Align seg 1/1 to: US-07-987-272A-14 from: 1 to: 89

```
1  AAGTTGACCGAGCTGAGAAAGCCTTGACTATCATGACGCTTACCA 50
|||||
1 MetProSerGlnLeuGlnLysAlaLeuSerAsnLeuIleAspValTyrH1 17
51 CAAGTACTCTCGATTAAGGGAATTTCCATGCTCTACAGGAGTAC 100
|||||
17 sAsnTyrSerAsnIleGlnGlnLysAsnHisAlaLeuTyrLysAsnAsp 34
101 TGAAGAAATTCCTAGACCGAGTGTCTCAGTATATCAGAAAAAGGT 150
|||||
34 heLysLysMetValThrThrGlnCysProGlnPheValGlnAsnIleAsn 50
151 GCAGACGCTGTGTTCAAGAGTTGGATATCAACACTGATGTGAGTTAA 200
|||||
51 IleGlnAsnLeuPheArgGlnLeuAspIleAsnSerAspAsnAlaIleAs 67
201 CTTCCAGAGTTCCTCATCTGTTGATTAAGATGGGCGTGCAGCCCA 250
|||||
67 nPheGlnGlnPheLeuAlaMetValIleLysValGlyAlaAlaSerHisL 84
251 AAAAAAGCCATGAAGAA 267
|||||
84 ySAspSerHisLysGln 89
```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-987-272A-1

## seq\_documentation\_block:

; Sequence 1, Application US/07987272A  
; Patent No. 5731166  
; GENERAL INFORMATION:  
; APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M  
; TITLE OF INVENTION: No. 5731166el Chemotactic Factor  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cushman Darby & Cushman  
; STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/987,272A  
; FILING DATE: 05-FEB-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PK 2127  
; FILING DATE: 05-FEB-1990

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 4463
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brinkman, David W
; REGISTRATION NUMBER: 20,817
; REFERENCE/DOCKET NUMBER: DWB/1925/200259
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861 3000
; TELEFAX: 202-822 0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-987-272A-1

```

```

alignment_scores:
  Quality: 293.00      Length: 87
  Ratio: 3.756         Gaps: 0
  Percent Similarity: 89.655   Percent Identity: 58.621

```

```

alignment_block:
US-09-806-382A-1 x US-07-987-272A-1 ..

```

Align seg 1/1 to: US-07-987-272A-1 from: 1 to: 88

```

7 ACCGAGCTGAGAGAAAGCCTTGAACTCTATCATCGACGTCTACCAAGTA 56
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
2 SerGluLeuGluLysAlaLeuSerAsnLeuIleAspValTyrHisAsnTy 18
57 CTCCTGATAAAGGGAATTTCCATGCCCTTACAGAGGATGACCTGACA 106
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 rSerAsnIleGlnGlyAsnHisAlaLeuTyrLysAsnAspPheLysL 35
107 AATTGCTAGACGACGAGTGTCTCAGTATATCAAGGAAAGGTCGACAGC 156
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35 ySmeValThrThrGluCysProGlnPheValGlnAsnIleAsnIleGlu 51
157 GTCTGCTCAAGAGCTGGATATCAACACTGATGGTGCAGTTAACTTCA 206
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52 AsnLeuPheArgGluLeuAspIleAsnSerAspAsnAlaIleAsnPheL 68
207 GCAGTTCCTCATCTGCGATTAAGATGGCGCTGCAGCCCAAAAAA 256
  :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 uGlnPheLeuAlaMetValIleLysValGlyAlaIleAsnHisLysAsp 85
257 GCCATGAGAA 267
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
85 erHisLysGlu 88

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-987-272A-17

```

seq_documentation_block:
; Sequence 17, Application US/07987272A
; Patent No. 5731166
; GENERAL INFORMATION:
; APPLICANT: Gecezy, C., Simpson, R. J. and Lackmann, M
; TITLE OF INVENTION: No. 5731166el Chemotactic Factor
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman Darby & Cushman
; STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/987, 272A
; FILING DATE: 05-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 2127
; FILING DATE: 05-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 4463
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brinkman, David W
; REGISTRATION NUMBER: 20,817
; REFERENCE/DOCKET NUMBER: DWB/1925/200259
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861 3000
; TELEFAX: 202-822 0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-987-272A-17

```

```

alignment_scores:
  Quality: 247.00      Length: 75
  Ratio: 3.687         Gaps: 0
  Percent Similarity: 89.333   Percent Identity: 57.333

```

alignment\_block:
US-09-806-382A-1 x US-07-987-272A-17 ..

Align seg 1/1 to: US-07-987-272A-17 from: 1 to: 76

```

7 ACCGAGCTGAGAGAAAGCCTTGAACTCTATCATCGACGTCTACCAAGTA 56
  ::::::::::::::::::::::::::::::::::::::::::::::::::::
2 SerGluLeuGluLysAlaLeuSerAsnLeuIleAspValTyrHisAsnTy 18
57 CTCCTGATAAAGGGAATTTCCATGCCCTTACAGAGGATGACCTGACA 106
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
18 rSerAsnIleGlnGlyAsnHisAlaLeuTyrLysAsnAspPheLysL 35
107 AATTGCTAGACGACGAGTGTCTCAGTATATCAAGGAAAGGTCGACAGC 156
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52 AsnLeuPheArgGluLeuAspIleAsnSerAspAsnAlaIleAsnPheL 68
157 GTCTGCTCAAGAGCTGGATATCAACACTGATGGTGCAGTTAACTTCA 206
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 GCAGTTCCTCATCTGCGATTAAG 231
  :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 uGlnPheLeuAlaMetValIleLys 76

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-200-016-2

```

seq_documentation_block:
; Sequence 2, Application US/08200016
; Patent No. 5614397
; GENERAL INFORMATION:
; APPLICANT: Weissman, Irving
; APPLICANT: Lagasse, Eric
; TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: APOPTOSIS IN HEMATOLYMPHOID CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Fish & Richardson  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: California  
COUNTRY: US  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/200,016  
FILING DATE: 22-FEB-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Rae-Venter, Barbara  
REGISTRATION NUMBER: 32,750  
REFERENCE/DOCKET NUMBER: 06037/003001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 854-5277  
TELEFAX: (415) 854-0875  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-200-016-2

alignment\_scores:  
Quality: 245.00 Length: 47  
Ratio: 5.213 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-806-382a-1 x US-08-200-016-2 ..

Align seg 1/1 to: US-08-200-016-2 from: 1 to: 47

1 ATGTTGACGAGCTGAGAAAGCTTGAACCTATCATGACGCTTACCA 50  
|||||  
1 MetleuThnGlueuGluuylsalaleuanserilelleaspvaltylhi 17  
17 slslytYrSerleuilellysglyAsnpheHisalavaltyrargaspaspl 34  
51 CAAGTACTCCTGATPAAAGGGAATTCCATGCCGTCTACAGGATGACC 100  
|||||  
101 TGAAGAAATTGCTAGAGCCGAGTCTCTCATATATGACG 141  
|||||  
34 euLysLysLeuLeuGluuTrgIucysProGlnTrpIlearg 47  
seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-200-016-3  
seq\_documentation\_block:  
Sequence 3, Application US/08200016  
Patent No. 5614397  
GENERAL INFORMATION:  
APPLICANT: Weissman, Irving  
APPLICANT: Lagasse, Eric  
TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR MODULATING  
TITLE OF INVENTION: APOPTOSIS IN HEMATOLYMPHOID CELLS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: California  
COUNTRY: US  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/200,016  
FILING DATE: 22-FEB-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Rae-Venter, Barbara  
REGISTRATION NUMBER: 32,750  
REFERENCE/DOCKET NUMBER: 06037/003001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 854-5277  
TELEFAX: (415) 854-0875  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 46 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-200-016-3

alignment\_scores:  
Quality: 240.00 Length: 46  
Ratio: 5.217 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-806-382a-1 x US-08-200-016-3 ..

Align seg 1/1 to: US-08-200-016-3 from: 1 to: 46

142 AAAAAGGTGACGAGCTGTGTTCAAGAGTTGATATCAACTGATGG 191  
|||||  
1 LyslysglyAlaaspvalTrpPhelysgluLeuaspIleasnThrpsgl 17  
192 TGCAGTTAACTCCAGAGATTCTTCATTTGGTGATPAAAGATGGCGTGG 241  
|||||  
17 yAlaValAsnpheGlnIupheLeuileValileLysmectlyVala 34  
242 CAGCCCAAAAAAGCCATGAGAAAGCCCAAGAG 279  
|||||  
34 laAlaHisLysLysSerHisGluLeuSerHisLysGlu 46  
seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-056-200-101  
seq\_documentation\_block:  
Sequence 101, Application US/08056200  
Patent No. 5616500  
GENERAL INFORMATION:  
APPLICANT: Steinert, Peter M.  
APPLICANT: Lee, Seung-Chul  
APPLICANT: Kim, In-Gyu  
APPLICANT: Chung, Soo-Il  
APPLICANT: Park, Sang-Chul  
TITLE OF INVENTION: Trichothyalin and Transglutaminase-3 and  
TITLE OF INVENTION: Methods of Using Same  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/056,200

```

; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-056-200-101

alignment_scores:
  Quality: 236.00      Length: 45
  Ratio: 5.244         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-806-382A-1 x US-08-056-200-101  ...
Align seg 1/1 to: US-08-056-200-101 from: 1 to: 45

7 ACCGAGCTGAGAAAGCCTTGAACTATATCAGCTACCAACAAGTA 56
|||||
1 ThnglueglulysAlaleuanserilleAspvaltyrlhsysty 17
|||||
57 CTCCTGATTAAGGGAATTTCCATCCGCTCTACAGGATGACCTGAAGA 106
|||||
17 rserleuilelysglyAsnphelhsAlavaltyrAtgAspAspleuLysL 34
|||||
107 AATGCTAGAGACCGAGTGCTCTCACTATATATCAGG 141
|||||
34 yslenuelugluthrglucysProglntlyrlleary 45

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-800-644-101

seq_documentation_block:
; Sequence 101, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997

```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/056,200
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-800-644-101

alignment_scores:
  Quality: 236.00      Length: 45
  Ratio: 5.244         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-806-382A-1 x US-08-800-644-101  ...
Align seg 1/1 to: US-08-800-644-101 from: 1 to: 45

7 ACCGAGCTGAGAAAGCCTTGAACTATATCAGCTACCAACAAGTA 56
|||||
1 ThnglueglulysAlaleuanserilleAspvaltyrlhsysty 17
|||||
57 CTCCTGATTAAGGGAATTTCCATCCGCTCTACAGGATGACCTGAAGA 106
|||||
17 rserleuilelysglyAsnphelhsAlavaltyrAtgAspAspleuLysL 34
|||||
107 AATGCTAGAGACCGAGTGCTCTCACTATATATCAGG 141
|||||
34 yslenuelugluthrglucysProglntlyrlleary 45

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-056-200-108

seq_documentation_block:
; Sequence 108, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```



```
APPLICATION NUMBER: US/08/056,200
FILING DATE: 30-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fedrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
US-08-056-200-108
```

```
alignment_scores:
Quality: 213.00      Length: 41
Ratio: 5.195         Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
US-09-806-382a-1 x US-08-056-200-108 ..
```

```
Align seg 1/1 to: US-08-056-200-108 from: 1 to: 41
```

```
142 AAAAAGGTGCAGAGCTGTGTTCAAGAGTTGATATCAACACTGATGG 191
|||||
1 LysLysGGLYAlaAspValTrpPheLysGluLeuAspIleAsnThrAspG1 17
192 TGCAGTTAACTTCCAGAGCTTCCTCATTTGTTGATTAAGATGGCGCTGG 241
|||||
17 YAlaValAsnPhelnglnuPheLeuIleLeuValIleLysMetGlyValA 34
242 CAGCCCAAAAAAAGCCATGAA 264
|||||
34 IaAlaHisLysLysSerHisGlu 41
```

```
seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-800-644-108
```

```
seq_documentation_block:
; Sequence 108, Application US/08800644
```

```
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
```

```
FILING DATE: 14-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/056,200
FILING DATE: 30-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fedrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
US-08-800-644-108
```

```
alignment_scores:
Quality: 213.00      Length: 41
Ratio: 5.195         Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000
```

```
alignment_block:
US-09-806-382a-1 x US-08-800-644-108 ..
```

```
Align seg 1/1 to: US-08-800-644-108 from: 1 to: 41
```

```
142 AAAAAGGTGCAGAGCTGTGTTCAAGAGTTGATATCAACACTGATGG 191
|||||
1 LysLysGGLYAlaAspValTrpPheLysGluLeuAspIleAsnThrAspG1 17
192 TGCAGTTAACTTCCAGAGCTTCCTCATTTGTTGATTAAGATGGCGCTGG 241
|||||
17 YAlaValAsnPhelnglnuPheLeuIleLeuValIleLysMetGlyValA 34
242 CAGCCCAAAAAAAGCCATGAA 264
|||||
34 IaAlaHisLysLysSerHisGlu 41
```

```
seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-568-310D-19
```

```
seq_documentation_block:
; Sequence 19, Application US/08568310D
```

```
; Patent No. 5976832
; GENERAL INFORMATION:
; APPLICANT: HITOMI, JIRO
; APPLICANT: YAMAGUCHI, KEN
; APPLICANT: YAMAMURA, TOKUJIRO
; APPLICANT: KIMURA, TATSUJI
; TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE
; STREET: 99 PARK AVENUE
; STREET: 6th FLOOR
; CITY: NEW YORK CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB
; COMPUTER: IBM-PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS 6.2
; SOFTWARE: WORDPERFECT 6.1
```

```

CORRESPONDENCE ADDRESS:
ADDRESSER: WYATT,GERRBER,MELLER & O'ROURKE
STREET: 99 PARK AVENUE
STREET: 6th FLOOR
CITY: NEW YORK CITY
STATE: NEW YORK
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE,3.50 INCH,720 Kb
MEDIUM TYPE: STORAGE
COMPUTER: IBM-PC COMPATIBLE
OPERATING SYSTEM: PC-DOS 6.2
SOFTWARE: WORDPERFECT 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/270,455
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/568,310
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KLEIN,MILTON
REGISTRATION NUMBER: 27101
REFERENCE/DOCKET NUMBER: 3316
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)953-3350
TELEFAX: (212)953-3352
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 92
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 19: FROM 1 TO 92
US-09-270-455-19

alignment_scores:
Quality: 160.00 Length: 92
Ratio: 2.500 Gaps: 1
Percent Similarity: 69.565 Percent Identity: 35.870

alignment_block:
US-09-806-382A-1 x US-09-270-455-19 ..

Align seg 1/1 to: US-09-270-455-19 from: 1 to: 92

      4 TTGACSCAGCTGGAGAAGGCSTTGACTGTATCATCGACSGTGAACA 53
        :::::::::::::::::::: ||:::|||||:||||:||||:
    1 MetThylusleuglnasrnlslseulgldgylelleasnllernlsgl 17
      54 GTTCGCCSCTTAAGGGAAATTTCCATGGCSGSCTCAAGGGATGASCTGA 10
        :|||:::||::|||:|||||:||||: |:::~::~|||
    17 nltyservvalargyalidlylsrneaprrhlleunslngsrfgeludell 34
      104 AGAAATGTCTAGAACSSCAGCTGCTCCAGTAGATAACAG..... 14
        ||:::||||: ||| ||::: |:||||:
    34 ysglneullethrplyglndeurproystrhpleglnaslntlrhspr 50
      142 AAAAAGGCTCACAGCTGTGGTTCAAAGACTGGATATCAACACTGATGG 19
        ::| | | | | | | | | | | | | | | | | | | | | |
    51 glmrptrhtlrlesrpilsrlrhneglnasrpluasrplsaspgrl 67
      192 TGGAGTTAACTTCAGAGGATTCCTCATTTGGTGATMAAGATGGCGGTG 24
        |||:||||:||||:||||:||||:||||:||||:||||:
    67 yaiaavalserrpheuldurhevaiallleualseratgyallleulyst 84
      242 CAGCCCACAAAAAAGCATGAAGA 267
        ::|||:||||| |:::|||||
    84 hrallahislslsapvllentlslvsqllu 92
```

```

seq_name: /cgn2_6/ptodata/2/iaa/SB_COMB.pep:US-08-568-310D-20
seq_documentation_block:
; Sequence 20, Application US/08568310D
; Patent No. 5976832
; GENERAL INFORMATION:
; APPLICANT: HITOMI, JIRO
; APPLICANT: YAMAGUCHI, KEN
; APPLICANT: YAMAMURA, TOKUJIRO
; APPLICANT: KIMURA, TATSUJI
; TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE
; STREET: 99 PARK AVENUE
; STREET: 6th FLOOR
; CITY: NEW YORK CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB
; MEDIUM TYPE: STORAGE
; COMPUTER: IBM-PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS 6.2
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,310D
; FILING DATE: DECEMBER 6, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7-70468 and 7-45564(both Japan)
; FILING DATE: 3/6/95 and 3/6/95, respectively
; ATTORNEY/AGENT INFORMATION:
; NAME: KLEIN, MILTON
; REGISTRATION NUMBER: 27101
; REFERENCE/DOCKET NUMBER: 3316
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)953-3352
; TELEFAX: (212)953-3350
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 20:
; RELEVANT RESIDUES IN SEQ ID NO: FROM 1 TO 92
; US-08-568-310D-20

alignment_scores:
      Quality: 158.00      Length: 96
      Ratio: 2.508
Percent Similarity: 65.625      Percent Identity: 38.542

alignment_block:
US-09-806-382a-1 x US-08-568-310D-20 ..

Align seg 1/1 to: US-08-568-310D-20 from: 1 to: 92

4 TTGACCGAGGTGGAGAAAGCCTGAACTGTATCATCGACGTCTACACAA 53
  ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|
1 MethrlyslengluglunlsleuglglYlleValasnllperhntsgl 17
  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
54 GTACTCCGTGATTAAGGGGAATTCCATGCCGTCTACAGGATGACCTGA 103
  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
17 nTysSerValargLysglYnlsPheasprnLleuSerLysglYgluLeuL 34
  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
104 AGAAATTGCTAGAGACCGAGTGTCTCAGTATATCAGG..... 141

```

```

||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
34 ysgInleuLeuThrLysGlulLeuAlaasnThrLleLysasnllLysasp 50
  ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
142 AAAAAGGCTGCAGACGTCGTGTTCAAGAGTTGATATATCAACACTGATG 191
  ||| ||| ||::|::|::|::|::|::|::|::|::|::|::|::|::|
51 LysAlaValllLeaspglulLephnglnglYleuaspAlaasnlnaspel 67
  ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
192 TGCAGTTAACTTCACGAGAGTTCCATTCGTGATGAAGATGGCGCTGG 241
  ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
67 uglInValaspPheglnnglupheLleSerLeuVal.....Alai 80
  ||| ||| ::::|::|::|::|::|::|::|::|::|::|::|::|::|
242 CACCCCACAAAAAAGCCATGAAGAAAGCCACAAAGAG 279
  ||| ||| ::::|::|::|::|::|::|::|::|::|::|::|::|::|
80 leaLaleuLysAlaAlaAlaHstYrnlstThrnlslYsglu 92

```

4

---

OM of: US-09-806-382a-1 to: PIR\_71:\* out\_format : pfs  
Date: Sep 9, 2002 3:12 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

# Command line parameters:

-MODEL=frame+\_r2p.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09806382/runtc\_09092002\_143848\_15421/app\_query.fasta\_1.752  
-DB=PIR\_71 -QFMT=fastan -SUFFIX=rpr -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOCL=0.000 -LOOEXT=0.000 -OGAPOP=4.500  
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELDP=6.000  
-DELXRT=7.000 -START=1 -MATRIX=blosome2 -TRANS=human40.cdt  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500  
-MINLEN=0 -MAXLEN=2000000000 -USER=US09806382.ecgn1.1.155  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPHY -WAIT -THREADS=1

## Search information block:

Query: US-09-806-382a-1  
Query length: 282  
Database: PIR\_71:\*  
Database sequences: 283138  
Database length: 96089334  
Search time (sec): 60.920000

seq_id	Strd Orig	ZScore	EScore	Len	Documentation
PIR1:JH0664	485.00	1040.00	8.3e-50	93	calgranulin A [validated] - human
PIR1:JH0665	320.00	684.67	5.4e-30	89	calgranulin A - rat
PIR1:156163	295.00	630.77	5.4e-27	89	calgranulin A - mouse
PIR2:B42628	170.00	368.36	4.9e-12	41	calcium-binding protein MRP-8 -
PIR2:BC4712	158.00	335.09	1.5e-10	92	S-100 protein beta chain [valida
PIR1:A442628	154.50	324.96	4.3e-10	122	calgranulin B - bovine (fragmen
PIR2:S24146	151.00	326.17	4.7e-10	95	S-100 protein P - human
PIR2:A55406	151.00	320.10	1.1e-09	91	calgranulin c - pig
PIR1:BCB01A	147.00	311.18	3.3e-09	94	S-100 protein alpha chain - bovi
PIR1:BCB01B	145.00	306.87	5.7e-09	94	S-100 protein alpha chain - huma
PIR2:A48015	139.00	294.13	3.0e-08	92	S-100 protein beta chain [valida
PIR2:A26557	138.00	291.97	3.9e-08	92	S-100 protein beta chain - mouse
PIR1:BCB01B	137.00	289.92	5.1e-08	91	S-100 protein beta chain - rat
PIR1:BCB01B	134.00	283.05	1.1e-07	95	S-100 protein alpha chain - bovi
PIR1:BCB01B	133.50	280.31	1.4e-07	114	calgranulin B [validated] - hun
PIR1:BCB01B	131.50	276.08	2.4e-07	113	calgranulin B - mouse
PIR1:BCB01B	130.50	273.92	3.2e-07	113	calgranulin B - rat
PIR2:A53219	126.00	265.25	1.1e-06	101	calvasculin - human
PIR2:A53217	125.00	263.18	1.4e-06	100	placental calcium-binding prote
PIR2:SO6207	124.00	260.93	1.9e-06	101	calvasculin - mouse
PIR1:Q01300	122.00	256.53	3.3e-06	102	calgizarin - rabbit
PIR2:SO1759	121.00	254.47	4.4e-06	101	calvasculin - rat
PIR2:A48319	120.50	252.61	5.1e-06	110	S-100 calcium-binding protein A
PIR1:BL001	119.50	253.48	6.3e-06	79	calcium-binding protein, intesti
PIR2:A41988	118.50	249.35	8.7e-06	98	S-100 calcium-binding protein A2
PIR1:BL001	115.50	244.86	1.9e-05	79	calcium-binding protein, intesti
PIR2:A45135	114.50	224.29	3.6e-05	591	proliferin - human (fragment)
PIR2:A48118	114.00	229.23	3.7e-05	306	major epidermal calcium-binding
PIR1:S27011	113.00	238.27	3.9e-05	90	calcylin - rabbit
PIR1:Q00246	112.50	238.39	4.4e-05	79	calcium-binding protein, intesti
PIR1:Q00603	112.00	236.51	4.5e-05	97	calpactin I light chain - chick
PIR2:A30129	112.00	235.25	5.2e-05	97	S-100 protein, lung - bovine
PIR2:S20342	111.00	235.25	5.2e-05	99	calcium-binding protein S100C -
PIR1:JH0780	111.00	232.55	7.0e-05	105	calgizarin - human
PIR1:BCB01B	110.00	231.81	9.0e-05	90	calcylin - human
PIR1:S28589	110.00	206.65	0.0001	1407	trichohyalin - rabbit
PIR1:A54327	109.50	229.67	0.0001	90	psoriasin [validated] - human
PIR2:B28363	109.00	229.65	0.0001	101	calcylin - rat
PIR2:A54314	108.00	227.60	0.0002	89	calcylin - mouse

seq\_name: PIR1:BCB01B  
seq\_documentation\_block:  
calgranulin A [validated] - human  
N:Alternate names: calcium-binding protein MRP-8; cystic fibrosis-associated antigen  
MRP-8; MIF-related protein 8k chain; oncodevelopmental protein; S-100 calcium-binding  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 07-Oct-1994 #text\_change 08-Dec-2000  
C:Accession: A31848; S00705; S13454; A29764; A60911; A61082; C54327; A44111; S53791  
R:Tagase, E.; Clerc, R.G.  
Mol. Cell. Biol. 8, 2402-2410, 1988  
A:Title: Cloning and expression of two human genes encoding calcium-binding proteins  
A:Reference number: A93102; MUID:86302148  
A:Accession: A31848  
A:Molecule type: DNA  
A:Residues: 1-93 <LAG>  
A:Cross-references: GB:M21005; NID:g188691; PIDN:AAA36327.1; PID:g386959  
R:Odink, K.; Cerletti, N.; Brueggemann, J.; Clerc, R.G.; Tarses, L.; Zwadlo, G.; Gerhar  
Nature 330, 80-82, 1987  
A:Title: Two calcium-binding proteins in infiltrate macrophages of rheumatoid arthritis  
A:Reference number: S00667; MUID:88039099  
A:Accession: S00705  
A:Molecule type: mRNA  
A:Residues: 1-93 <ODI>  
A:Cross-references: EMBL:X06234; NID:g34772; PIDN:CAA29580.1; PID:g34773  
A:Note: Part of this sequence, including the amino end of the mature protein, was con  
R:Schaefer, T.; Sachse, G.E.; Gassen, H.G.  
Biochem. Biophys. Res. Commun. 172, 1-4, 1991  
A:Title: The calcium-binding protein MRP-8 is produced by human pulmonary tumor cells  
A:Reference number: S13454; MUID:91248411  
A:Accession: S13454  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-93 <SCH>  
R:Dorin, J.R.; Novak, M.; Hill, R.E.; Brock, D.J.H.; Secher, D.S.; van Heyningen, V.  
Nature 326, 614-617, 1987  
A:Title: A clue to the basic defect in cystic fibrosis from cloning the CF antigen ge  
A:Reference number: A29764; MUID:87173041  
A:Accession: A29764  
A:Molecule type: mRNA  
A:Residues: 1-78, 'AMPTKAKKAKKRS' <DOR>  
A:Cross-references: EMBL:Y00278; NID:g29887; PIDN:CAA68390.1; PID:g29888  
A:Note: the differences after residue 78 are due to a frameshift  
R:Andersen, K.B.; Sletten, K.; Bernitz, H.B.; Dale, I.; Brandtzaeg, P.; Jellum, E.;  
Scand. J. Immunol. 28, 241-245, 1988  
A:Title: The leucocyte LI protein: identity with the cystic fibrosis antigen and the  
A:Reference number: A60911; MUID:88321575  
A:Accession: A60911  
A:Molecule type: protein  
A:Residues: 1-53, 'X', 55-56, 'X', 58-59, 78-84, 'X', 86-89 <AND>  
R:Tobe, T.; Murakami, K.; Tomita, M.; Nozawa, R.  
Chem. Pharm. Bull. 37, 1576-1580, 1989  
A:Title: Amino acid sequences of 608B antigens induced in HU-60 cells by 1,25-dihydro  
A:Reference number: A61082; MUID:89376638  
A:Accession: A61082  
A:Molecule type: protein  
A:Residues: 1-27, 29-33, 40-46, 58-69, 71-81, 'S', 83-84, 'S' <TOB>  
A:Experimental source: chronic myeloid leukemic cell line  
R:Madson, P.; Rasmussen, H.H.; Leffers, H.; Honore, B.; Dejgaard, K.; Olsen, E.; Kill  
E.  
Invest. Dermatol. 97, 701-712, 1991  
A:Title: Molecular cloning, occurrence, and expression of a novel partially secreted  
A:Accession: C54327  
A:Molecule type: protein  
A:Residues: 'X', 38-47, 50-53, 'X', 55-56 <MAD>  
R:Lemarchand, P.; Vaglio, M.; Manuel, J.; Markert, M.

J. Biol. Chem. 267, 19379-19382, 1992  
A:Title: Translocation of a small cytosolic calcium-binding protein (MRP-8) to plasma membrane  
A:Reference number: A44111; MUID:92406885  
A:Accession: A44111  
A:Molecule type: protein  
A:Residues: 1-25 <LEMB>  
A:Experimental source: zymosan-activated neutrophils  
A:Note: sequence extracted from NCBI backbone (NCBI:P113895)  
R:Nakai, M.; Ishikawa, M.; Hamada, Y.; Sugano, S.  
Biol. Chem. Hoppe-Seyler 375, 789-792, 1994  
A:Title: Isolation of an ascitic oncodevelopmental protein exhibiting high sequence homology  
A:Reference number: S53791; MUID:95209785  
A:Accession: S53791  
A:Molecule type: protein  
A:Residues: 1-16, 'X', '18-20 <NAN>  
C:Comment: Concentrations of this protein, a product of normal and leukemic granulocytes  
C:Comment: This protein can bind two calcium ions per molecule with an affinity similar  
C:Genetics:  
A:Gene: GDB:S100A8; CAGA; CPAG  
A:Cross-references: GDB:120569; OMIM:123885  
A:Map position: 1q21-1q21  
A:Introns: 47/73  
A:Note: the first intron occurs before the initiator codon  
C:Complex: homodimer; heterodimer and higher complexes with calgranulin B (see PIR:B3184)  
C:Superfamily: S-100 protein; calmodulin repeat homology  
C:Keywords: calcium binding; cytokine; EF hand; heterodimer; homodimer; inflammation  
F:1-93/Product: calgranulin A status experimental <MAT>  
F:7-41/Domain: calmodulin repeat homology <EP1>  
F:46-78/Domain: calmodulin repeat homology <EP2>

alignment\_scores:  
Quality: 485.00 Length: 93  
Ratio: 5.215 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-806-382A-1 x BCHUCF ..

Align seg 1/1 to: BCHUCF from: 1 to: 93

```
1  ATGTGACCGAGCTGGAGAAAGCCTGATCATCGAGCTACCA 50
|||||
1  MetLeuThrGluLeuGluLysAlaLeuAsnSerIleIleAspValTyrH 17
51  CAAGTACTCCCTGATTAAGGGAATTCATCCGCTACAGGAGTACC 100
|||||
17  slystYrSerIleuIleLysGlyAsnPhenIleAlaValTyrArgAspAsp 34
34  eulYsLysLeuLeuGluThrGluCysProGlnTyrIleArgLysLysGly 50
101  TGAAGAAATTCCTAGAGACCGAGTGTCTCATATATCAGAAAAAGGT 150
|||||
34  eulYsLysLeuLeuGluThrGluCysProGlnTyrIleArgLysLysGly 50
151  GCAGACGCTGCTGTTCAAGAGTTGGATATCACACGATGGTGCAGTTAA 200
|||||
51  AlaAspValTyrPheLysGluLeuAspIleAsnThrAspGlyAlaValAs 67
201  CTTCAGAGTTCCTCATCTGTTGATTAAGATGGCGTGCAGCCACACA 250
|||||
67  nPhenGlnGluPheLeuIleLeuValIleLysMetGlyAlaIleAlaHisL 84
251  AAAAAAGCCATGAGAAAGCCAAAGAG 279
|||||
84  yLysSerHisGluLeuSerHisLysGlu 93
```

seq\_name: p1r1:JN0685

seq\_documentation\_block:

calgranulin A - rat  
N:Alternate names: calcium-binding protein MRP-8; macrophage migration inhibitory factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 03-May-1994 #sequence\_revision 23-May-1997 #text\_change 22-Jun-1999  
C:Accession: JN0685

R:Imamiuchi, T.; Uchida, I.; Wahl, S.M.; McCartney-Francis, N.  
Biochem. Biophys. Res. Commun. 194, 819-825, 1993  
A:Title: Expression and cloning of migration inhibitory factor-related protein (MRP8)  
A:Reference number: JN0685; MUID:93343942  
A:Accession: JN0685  
A:Molecule type: mRNA  
A:Residues: 1-89 <IMA>  
A:Cross-references: GB:LI8891; NID:g349548; PIDN:AAA41637.1; PID:g349549  
A:Experimental source: strains LEW/N and F344/N  
C:Comment: This protein has a role in susceptibility to SCW-induced chronic disease.  
C:Genetics:  
A:Gene: MRP8  
C:Complex: homodimer; heterodimer and higher complexes with calgranulin B  
C:Superfamily: S-100 protein; calmodulin repeat homology  
C:Keywords: calcium binding; cytokine; EF hand; heterodimer; homodimer; inflammation  
F:7-41/Domain: calmodulin repeat homology <EP1>  
F:46-78/Domain: calmodulin repeat homology <EP2>

alignment\_scores:  
Quality: 320.00 Length: 89  
Ratio: 3.951 Gaps: 0  
Percent Similarity: 91.011 Percent Identity: 62.921

alignment\_block:

US-09-806-382A-1 x JN0685 ..

Align seg 1/1 to: JN0685 from: 1 to: 89

```
1  ATGTGACCGAGCTGGAGAAAGCCTGATCATCGAGCTACCA 50
|||||
1  MetLeuThrGluLeuGluLysAlaLeuAsnSerIleIleGluValTyrH 17
51  CAAGTACTCCCTGATTAAGGGAATTCATCCGCTACAGGAGTACC 100
|||||
17  sasnYrSerGlyIleLysGlyAsnHisAlaLeuTyrArgAspAsp 34
34  heargLysMetValThrThrGluCysProGlnPheValGlnAsnLysAsn 50
101  TGAAGAAATTCCTAGAGACCGAGTGTCTCATATATCAGAAAAAGGT 150
|||||
34  heargLysMetValThrThrGluCysProGlnPheValGlnAsnLysAsn 50
151  GCAGACGCTGCTGTTCAAGAGTTGGATATCACACGATGGTGCAGTTAA 200
|||||
51  ThrGluSerLeuPheLysGluLeuAspValAsnSerAspAsnAlaIleAs 67
201  CTTCAGAGTTCCTCATCTGTTGATTAAGATGGCGTGCAGCCACACA 250
|||||
67  nPhenGlnGluPheLeuAlaLeuValIleArgValGlyAlaIleAlaHisL 84
251  AAAAAAGCCATGAGAA 267
|||
84  yAspSerHisLysGlu 89
```

seq\_name: p1r1:I56163

seq\_documentation\_block:

calgranulin A - mouse  
N:Alternate names: calcium-binding protein MRP-8; CP-10 chemotactic protein; macrophage  
C:Species: Mus musculus (house mouse)  
C:Date: 26-Jul-1996 #sequence\_revision 23-May-1997 #text\_change 22-Jun-1999  
C:Accession: I56163; A42488  
R:Lackmann, M.; Rajasekariah, P.; Iismaa, S.E.; Jones, G.; Cornish, C.J.; Hu, S.; Slin  
J. Immunol. 150, 2981-2991, 1993  
A:Title: Identification of a chemotactic domain of the pro-inflammatory S100 protein  
A:Reference number: I56163; MUID:93203618  
A:Accession: I56163  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-89 <LACI>  
A:Cross-references: GB:S57123; NID:g298706; PIDN:AA825840.1; PID:g298707  
R:Lackmann, M.; Cornish, C.J.; Simpson, R.J.; Moritz, R.L.; Geczy, C.L.  
J. Biol. Chem. 267, 7499-7504, 1992  
A:Title: Purification and structural analysis of a murine chemotactic cytokine (CP-10

A:Reference number: A42488; MUID:92218405  
 A:Accession: A42488  
 A:Molecule type: protein  
 A:Residues: 2-77 <LAC2>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:94068)  
 C:Genetics: MRP8  
 C:Complex: homodimer; heterodimer and higher complexes with calgranulin B  
 C:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: calcium binding; cytokine; EF hand; heterodimer; inflammation  
 F:2-89/Product: calgranulin A #status experimental <MAT>  
 F:7-41/Domain: calmodulin repeat homology <EF1>  
 F:46-78/Domain: calmodulin repeat homology <EF2>

alignment\_scores:  
 Quality: 295.00 Length: 89  
 Ratio: 3.734 Gaps: 0  
 Percent Similarity: 88.764 Percent Identity: 58.427

alignment\_block:  
 US-09-806-382a-1 x I56163 ..

Align seg 1/1 to: I56163 from: 1 to: 89

```

1  ATGTGACCGAGCTGGAGAAACCTTGAACTATCATCGACGCTTACCA 50
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1  MetProSerGluLeuGluLysAlaLeuSerAsnLeuIleAspValTyrH1 17
51  CAAGTACCTCCCTGATTAAGGGAATTTCCATGCGCTACAGGATGACC 100
   |||:||||| |||:||||| |||:|||||:|||||:|||||:|||||:|||||
17  ssnrYrSerAsnIleGlnGlyAsnHisAlaLeuTyrLysAsnAspP 34
101  TGAAGAAATTGCTAGAGACCGAGTCTCTCAGTATATCAGAAAAAGGT 150
   :|||:|||||: |||:|||||:|||||:|||||:|||||: |||:
34  helysysMetValThrThrGluCysProGlnPheValGlnAsnIleAsn 50
151  GCAGACGTGTGTTCAAGAGTTGATATCAACACTGATGTCGACGTTAA 200
   :|||: |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51  IleGlnAsnLeuPheArgGluLeuAspIleAsnSerAspAsnAlaIleAs 67
201  CTTCACGAGATCTCATCTGCTGATTAAGATGGCGGACGACCCACA 250
   |||:|||||:|||||: |||:|||||:|||||:|||||:|||||:|||||
67  nPheGluGluPheLeuAlaMetValIleLysValGlyAlaIleSerHisL 84
251  AAAAAAGCCATGAAGAA 267
   || |||:|||||:|||||
84  ysAspSerHisLysGlu 89

```

seq\_name: p1r2:B42628

seq\_documentation\_block:  
 calcium-binding protein MRP-8 - bovine (fragment)  
 N:Alternate names: neutrophil cytosolic 7K phosphoprotein; nuclear protein 1  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 29-Sep-1999  
 C:Accession: B42628; A22309  
 R:Dianoux, A.C.; Stasia, M.J.; Garin, J.; Gagnon, J.; Vignais, P.V.  
 Biochem. J. 31, 589S-590S, 1992  
 A:Title: The 23-kDa cytosolic protein, a substrate of protein kinase C, in bovine neutrophils  
 A:Reference number: A42628; MUID:92304974  
 A:Accession: B42628  
 A:Molecule type: protein  
 A:Residues: 1-41 <DIA>  
 R:Tang, T.K.; Hong, T.M.; Lin, C.Y.; Lai, M.L.; Liu, C.H.L.; Lo, H.J.; Wang, M.E.; Chen,  
 submitted to the Protein Sequence Database, July 1992  
 A:Reference number: A22309  
 A:Accession: A22309  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-9, 'NI', '12-21', 'K', '23-29' <TAN>  
 C:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: calcium binding; cytosol; EF hand; phosphoprotein

alignment\_scores:  
 Quality: 170.00 Length: 41  
 Ratio: 4.595 Gaps: 0  
 Percent Similarity: 90.244 Percent Identity: 78.049

alignment\_block:  
 US-09-806-382a-1 x B42628 ..

Align seg 1/1 to: B42628 from: 1 to: 41

```

1  ATGTGACCGAGCTGGAGAAACCTTGAACTATCATCGACGCTTACCA 50
   |||:||||| |||:||||| |||:|||||:|||||:|||||:|||||:|||||
1  MetLeuThrAspLeuGlu**AlaIleAspSerLeuIleAspValTyrH1 17
51  CAAGTACCTCCCTGATTAAGGGAATTTCCATGCGCTACAGGATGACC 100
   |||:||||| |||:||||| |||:|||||:|||||:|||||:|||||:|||||
17  slystYrSerLeu**LysGlyAsnTyrHisAlaValTyr**AspAspL 34
101  TGAAGAAATTGCTAGAGACCGAG 123
   ||| |||:|||||:|||||
34  eulys***LeuLeuGluThrGlu 41

```

seq\_name: p1r2:JC4712

seq\_documentation\_block:  
 S-100 calcium-binding protein A12 - human  
 N:Alternate names: calcium-binding amniotic fluid protein 1 (CAAF1); calgranulin C; c  
 utrophil protein  
 C:Species: Homo sapiens (man)  
 C:Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 13-Aug-1999  
 C:Accession: JC4712; JC4717; JC4891; S56113; S56114  
 R:Yamamura, T.; Hiltomi, J.; Nagasaki, K.; Suzuki, M.; Takahashi, E.; Saito, S.; Tsuka  
 Biochem. Biophys. Res. Commun. 221, 356-360, 1996  
 A:Title: Human CAAF1 gene - molecular cloning, gene structure, and chromosome mapping  
 A:Reference number: JC4712; MUID:96192053  
 A:Accession: JC4712  
 A:Molecule type: mRNA  
 A:Residues: 1-92 <YAM>  
 A:Cross-references: DDBJ:D83657; NID:g1502284; PIDN:BA12030.1; PID:g1502285  
 R:Marit, T.; Ertmann, K.D.; Gallin, M.Y.  
 Biochem. Biophys. Res. Commun. 221, 454-458, 1996  
 A:Title: Host-parasite interaction in human onchocerciasis: Identification and sequen  
 A:Reference number: JC4717; MUID:96192069  
 A:Accession: JC4717  
 A:Molecule type: protein  
 A:Residues: 2-92 <MAR>  
 A:Experimental source: Onchocerca volvulus infecting human tissue  
 R:Ilig, E.C.; Troxler, H.; Bueglisser, D.M.; Kuster, T.; Markert, M.; Guignard, F.; Hu  
 Biochem. Biophys. Res. Commun. 225, 146-150, 1996  
 A:Title: Amino acid sequence determination of human S100 A12 (P6, Calgranulin C, CGRP  
 A:Reference number: JC4891; MUID:96332419  
 A:Accession: JC4891  
 A:Molecule type: protein  
 A:Residues: 2-92 <ILG>  
 R:Guignard, F.; Maue, J.; Markert, M.  
 Biochem. J. 309, 395-401, 1995  
 A:Title: Identification and characterization of a novel human neutrophil protein rela  
 A:Reference number: S56113; MUID:95351965  
 A:Accession: S56113  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 'XX', '4-14', 'X', '16-17', 'XXXX' <GUT1>  
 A:Experimental source: isoform 6a  
 A:Accession: S56114  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 2-21 <GUT2>  
 A:Experimental source: isoform 6b  
 C:Comment: This protein is released by activated neutrophils in the course of inflam  
 C:Genetics:  
 A:Gene: GDB:S100A12; p6; MRP6; CGRP; CAAF1

A:Cross-references: GDB:5218374  
 A:Map position: 1q21-1q21  
 C:Complex: monomer  
 C:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: calcium binding; EF hand; monomer; neutrophil; zinc  
 F:2-92/Product: S-100 calcium-binding protein A12 #status experimental <MAT>  
 F:6-39/Domain: calmodulin repeat homology <EF1>  
 F:49-81/Domain: calmodulin repeat homology <EF2>  
 F:86-90/Region: zinc binding #status predicted

alignment\_scores:  
 Quality: 158.00 Length: 96  
 Ratio: 2.508 Gaps: 2  
 Percent Similarity: 65.625 Percent Identity: 38.542

alignment\_block:  
 US-09-806-382A-1 x JC4712 ..

Align seg 1/1 to: JC4712 from: 1 to: 92

```

4 TTGACCGAGCTGGAGAAAGCCTTGAACCTATCATGACGCTACCAAA 53
  ::::::::::::::::::::|::::::::::::::::::|:::
1 MetThrLysLeuGlnGluHisLeuGlnGlyIleValAsnIlePheHisG1 17
54 GTACTCCCTGATTAAGGGGAATTTCCATGCCGTACAGGATGACCTGA 103
  ::::::::::::::::::::|::::::::::::::::::|:::
17 nHySerValArgLysGlyHisPheAspThrLeuSerLysGlyGluLeu 34
104 AGAAATTCGTAGAGACCGAGTGTCCCTCACTATATCAGG..... 141
  ::::::::::::::::::::|::::::::::::::::::|:::
34 yGlnLeuLeuThrLysGlnLeuAlaAsnThrIleLysAsnIleLysAsp 50
142 AAAAAGGGTGCAGACGCTGTGTTCAAGATGATATCAACACTATGG 191
  ::::::::::::::::::::|::::::::::::::::::|:::
51 LysAlaValIleAspGlnIlePheGlnGlyLeuAspAlaAsnGlnAspG1 67
192 TGCAGTTAACTTCCAGAGAGTTCCTCATTTCTGGATTAAGATGGCGGTG 241
  ::::::::::::::::::::|::::::::::::::::::|:::
67 uGlnValAspPheGlnGluPheIleSerLeuVal.....AlaI 80
242 CAGCCCAACAAAAAGCCATGAAGAAAGCCACAAAGAG 279
  ::::::::::::::::::::|::::::::::::::::::|:::
80 LeuLeuLysAlaAlaHisIleThrHisThrHisLysGln 92

```

seq\_name: pIrl:A42628

seq\_documentation\_block:

N:Alternate names: calcium-binding protein MRP-14; macrophage migration inhibitory factor in 2  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Sep-1993 #sequence\_revision 23-May-1997 #text\_change 23-May-1997  
 C:Accession: B22309; A42628  
 R:Yang, T.K.; Hong, T.M.; Lin, C.Y.; Lai, M.L.; Liu, C.H.L.; Lo, H.J.; Wang, M.E.; Chen, Submitted to the Protein Sequence Database, July 1992  
 A:Reference number: A22309  
 A:Accession: B22309  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-122 <TAN>  
 R:Dianoux, A.C.; Stasia, M.J.; Garlin, J.; Gagnon, J.; Vignais, P.V.  
 Biochemistry 31, 5898-5905, 1992  
 A:Title: The 23-kilodalton protein, a substrate of protein kinase C, in bovine neutrophil  
 A:Reference number: A42628; MUID:92304974  
 A:Accession: A42628  
 A:Molecule type: protein  
 A:Residues: 4-32, 'F', 34-56 <DIA>  
 C:Complex: heterodimer and higher complexes with calgranulin A  
 C:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: blocked amino end; calcium binding; EF hand; heterodimer; inflammation; phosphatase  
 F:6-40/Domain: calmodulin repeat homology <EF1>  
 F:50-82/Domain: calmodulin repeat homology <EF2>

alignment\_scores:  
 Quality: 154.50 Length: 91  
 Ratio: 2.239 Gaps: 1  
 Percent Similarity: 75.824 Percent Identity: 27.473

alignment\_block:  
 US-09-806-382A-1 x A42628 ..

Align seg 1/1 to: A42628 from: 1 to: 122

```

4 TTGACCGAGCTGGAGAAAGCCTTGAACCTATCATGACGCTACCAAA 53
  ::::::::::::::::::::|::::::::::::::::::|:::
1 MetSerGlnMetGlnSerSerIleGlnThrIleIleAsnIlePheHisG1 17
54 GTACTCCCTGATTAAGGGGAATTTCCATGCCGTACAGGATGACCTGA 103
  ::::::::::::::::::::|::::::::::::::::::|:::
17 nHySerValArgLysGlyHisPheAspThrLeuSerLysGlyGluLeu 34
104 AGAAATTCGTAGAGACCGAGTGTCCCTCACTATATCAGGAAAG..... 147
  ::::::::::::::::::::|::::::::::::::::::|:::
34 yGlnLeuValGlnLysGlnLeuProAsnPheLeuLysGlnLysLys 50
51 AsnGlnAlaAlaIleAsnGlnIleMetGlnAspLeuAspThrAsnValas 67
148 .....GTGCAGACGCTGTGTTCAAGATGATATCAACACTGA 188
  ::::::::::::::::::::|::::::::::::::::::|:::
189 TGTGCGAGTTAACTTCCAGAGAGTTCCTCATTTCTGGATTAAGATGGCG 238
  ::::::::::::::::::::|::::::::::::::::::|:::
67 pLysGlnLeuSerPheGlnGluPheIleMetGlnAspLeuValAlaArgLeuThr 84
239 TGCAGCCCAACAAAAAGCCAT 261
  ::::::::::::::::::::|::::::::::::::::::|:::
84 AlaIleSerHisGlnGluMetHis 91

```

seq\_name: pIrl:S24146

seq\_documentation\_block:

S-100 protein p - human  
 C:Species: Homo sapiens (man)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Aug-1999  
 C:Accession: S24146; PS0340  
 R:Becker, T.; Gerke, V.; Kube, E.; Weber, K.  
 Eur. J. Biochem. 207, 541-547, 1992  
 A:Title: S100P, a novel Ca(2+)-binding protein from human placenta. cDNA cloning, rec  
 A:Reference number: S24146; MUID:92339442  
 A:Accession: S24146  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-95 <BEC>  
 A:Cross-references: EMBL:X65614; NID:g36177; PIDN:CAA46566.1; PID:g36178  
 R:Emoto, Y.; Kobayashi, R.; Akatsuka, H.; Hidaka, H.  
 Biochem. Biophys. Res. Commun. 182, 1246-1253, 1992  
 A:Title: Purification and characterization of a new member of the S-100 protein family  
 A:Reference number: PS0340; MUID:92171935  
 A:Accession: PS0340  
 A:Molecule type: protein  
 A:Residues: 1-31, 'F', 33-84, 'X', 86-91 <EMO>  
 A:Experimental source: placenta  
 C:Genetics:  
 A:Gene: GDB:S100P  
 A:Cross-references: GDB:134405; OMIM:600614  
 A:Map position: 4p16-4p16  
 C:Superfamily: S-100 protein; calmodulin repeat homology  
 C:Keywords: calcium binding; EF hand; placenta  
 F:6-40/Domain: calmodulin repeat homology <EF1>  
 F:49-81/Domain: calmodulin repeat homology <EF2>

alignment\_scores:  
 Quality: 154.00 Length: 87  
 Ratio: 2.525 Gaps: 1







US-09-806-382a-1 x BCHUIB ..

Align seg 1/1 to: BCHUIB from: 1 to: 92

```

4 TTGACCGAGCTGAGAAAGCCTTGAATCTATCATCGACGCTTACCACAA 53
   :::::::::::::::::::::
1 MetSerGluLeuGluValMetValAlaLeuIleAspValPheHisG1 17
54 GTCATCCCGATAAAGGGAATTTCCATGCCGCTACAGGAGTGCCTGA 103
   :::::::::::::::::::::
17 nlyrSerGlyArgGluGlyAspLysHisLysLysLysSerGluLeuL 34
104 AGAAATTCGTAGAGACCGAGTCTCCTCATGTATATC.....AGG 141
   :::::::::::::::::::::
34 ysgluleuIleasnsgluleuSerHisPheLeuGluGluIleLysGlu 50
142 AAAAAGGTGACAGACGCTGTGTTCAAGAATTGGATATCAACACTGATG 191
   :::::::::::::::::::::
51 GIngluValAlaAspLysValMetGluThrIleuAspAsnAspGlyAspG1 67
192 TGCAGTTAACTCCAGAGCTTCCTCATCTGTGTCGTTAAAGATGGCGCTG 241
   :::::::::::::::::::::
67 yglucysaspheGlnGluPheMetAlaPheValAlaMetValThrThra 84
242 CAGCCACACAA 252
   :::::::::::::::::::::
84 lAcysHisGlu 87

```

seq\_name: pir2:A48015

seq\_documentation\_block:

S-100 protein beta chain - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 24-Nov-1999  
C:Accession: A48015  
R:Jiang, H.; Shah, S.; Hilt, D.C.  
J. Biol. Chem. 268, 20502-20511, 1993  
A:Title: Organization, sequence, and expression of the murine S100beta gene. Transcript  
A:Reference number: A48015; MUID:93388628  
A:Accession: A48015  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-92 <JIA>  
A:Cross-references: GB:I22144; NID:g404768; PIDN:AAA03075.1; PID:g404769  
C:Genetics:  
A:Introns: 46/3  
C:Superfamily: S-100 protein; calmodulin repeat homology  
C:Keywords: acetylated amino end; calcium binding; EF hand  
F:2-92/Product: S-100 protein beta chain #status predicted <MAT>  
F:6-40/Domain: calmodulin repeat homology <EF1>  
F:49-81/Domain: calmodulin repeat homology <EF2>  
F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted  
F:20,22,24,27,32/Binding site: calcium (Gly, Glu, Asp, Lys, Glu) #status predicted  
F:62,64,66,68,70,73/Binding site: calcium (Asp, Asp, Glu, Asp, Glu) #status predicted

alignment\_scores:

Quality:	Length:
139.00	87
Ratio: 2.206	Gaps: 1
Percent Similarity: 72.414	Percent Identity: 33.333

alignment\_block:

US-09-806-382a-1 x A48015 ..

Align seg 1/1 to: A48015 from: 1 to: 92

```

4 TTGACCGAGCTGAGAAAGCCTTGAATCTATCATCGACGCTTACCACAA 53
   :::::::::::::::::::::
1 MetSerGluLeuGluValMetValAlaLeuIleAspValPheHisG1 17
54 GTCATCCCGATAAAGGGAATTTCCATGCCGCTACAGGAGTGCCTGA 103
   :::::::::::::::::::::
17 nlyrSerGlyArgGluGlyAspLysHisLysLysLysSerGluLeuL 34

```

```

104 AGAAATTCGTAGAGACCGAGTCTCCTCATGTATATC.....AGG 141
   :::::::::::::::::::::
34 ysgluleuIleasnsgluleuSerHisPheLeuGluGluIleLysGlu 50
142 AAAAAGGTGACAGACGCTGTGTTCAAGAATTGGATATCAACACTGATG 191
   :::::::::::::::::::::
51 GIngluValAlaAspLysValMetGluThrIleuAspAsnAspGlyAspG1 67
192 TGCAGTTAACTCCAGAGCTTCCTCATCTGTGTCGTTAAAGATGGCGCTG 241
   :::::::::::::::::::::
67 yglucysaspheGlnGluPheMetAlaPheValAlaMetValThrThra 84
242 CAGCCACACAA 252
   :::::::::::::::::::::
84 lAcysHisGlu 87

```

seq\_name: pir2:A26557

seq\_documentation\_block:

S-100 protein beta chain - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 25-Mar-1988 #sequence\_revision 04-Nov-1994 #text\_change 13-Aug-1999  
C:Accession: A60046; S07357; A26557  
R:Maeda, T.; Usui, H.; Araki, K.; Kuwano, R.; Takahashi, Y.; Suzuki, Y.  
Brain Res. Mol. Brain Res. 10, 193-202, 1991  
A:Title: Structure and expression of rat S-100 beta subunit gene.  
A:Reference number: A60046; MUID:91359841  
A:Accession: A60046  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-92 <MAE>  
A:Cross-references: GB:S53527  
R:Kuwano, R.; Usui, H.; Maeda, T.; Fukui, T.; Yamanari, N.; Ohtsuka, E.; Ikehara, M.;  
Nucleic Acids Res. 12, 7455-7465, 1984  
A:Title: Molecular cloning and the complete nucleotide sequence of cDNA to mRNA for S  
A:Reference number: S07357; MUID:85037524  
A:Accession: S07357  
A:Molecule type: mRNA  
A:Residues: 1-92 <KUM>  
A:Cross-references: EMBL:X01090; NID:957174; PIDN:CAA25567.1; PID:957175  
R:Dunn, R.; Landry, C.; O'Hanlon, D.; Dunn, J.; Allore, R.; Brown, I.; Marks, A.  
J. Biol. Chem. 262, 3562-3566, 1987  
A:Title: Reduction in S100 protein beta-subunit mRNA in C6 rat glioma cells following  
A:Reference number: A26557; MUID:87137648  
A:Accession: A26557  
A:Molecule type: mRNA  
A:Residues: 6-92 <DUN>  
A:Cross-references: GB:M15705  
C:Comment: S-100 protein occurs as alpha-beta heterodimers, alpha-alpha heterodimers,  
C:Genetics:  
A:Introns: 46/3  
C:Superfamily: S-100 protein; calmodulin repeat homology  
C:Keywords: brain; calcium binding; dimer; EF hand; zinc  
F:2-92/Product: S-100 protein beta chain #status predicted <MAT>  
F:6-40/Domain: calmodulin repeat homology <EF1>

alignment\_scores:

Quality:	Length:
138.00	87
Ratio: 2.190	Gaps: 1
Percent Similarity: 72.414	Percent Identity: 33.333

alignment\_block:

US-09-806-382a-1 x A26557 ..

Align seg 1/1 to: A26557 from: 1 to: 92

```

4 TTGACCGAGCTGAGAAAGCCTTGAATCTATCATCGACGCTTACCACAA 53
   :::::::::::::::::::::
1 MetSerGluLeuGluValMetValAlaLeuIleAspValPheHisG1 17
54 GTCATCCCGATAAAGGGAATTTCCATGCCGCTACAGGAGTGCCTGA 103

```

```
||||| ..... ||| ... :|||||
17 nYrSerGIyArggLuGIyAspLysHIsLysLeuLysSerGIuLeuL 34
104 AGAAATGCTAGAGACCGAGTCTCTCACTATATC.....AGG 141
34 YSLuLeuLLeAsnSngLuLeuSerHIsPheLeuLuGIuLeuLysGIu 50
142 AAAAAGGGTCAGAGCTGTGGTTCAAGAGTGGATATCAACAGCTGATG 191
51 GInGIuValAlAspLysValMetGIuThrLeuAspGIuAspGIu 67
192 TGCAGTTAACTTCAGAGATCTCTCTTCTGGTGATMAAGATGGCGGTG 241
67 YGIuLysAspPheGInGIuPheMetAlaPheValSerMetValThrThra 84
242 CAGCCACAA 252
84 LacYSHISGIu 87

seq_name: pIrl:BCBOIB

seq_documentation_block:
S-100 protein beta chain - bovine
N:Alternate names: neurocalcin delta-binding protein S100-beta
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-May-1979 #sequence_revision 14-Nov-1983 #text_change 24-Nov-1999
C:Accession: A91254; B91110; A90075; S54348; A03077
R:Isobe, T.; Okuyama, T.
Eur. J. Biochem. 89, 379-388, 1978
A:Title: The amino-acid sequence of S-100 protein (PAP-I-b protein) and its relation to
A:Reference number: A91254; MUID:79045265
A:Accession: A91254
A:Molecule type: protein
A:Residues: 'ESEL', 5-91 <ISO>
A:Experimental source: brain
A:Note: This sequence has since been revised in reference A91110
R:Isobe, T.; Okuyama, T.
Eur. J. Biochem. 116, 79-86, 1981
A:Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a protein.
A:Reference number: A91110; MUID:81236562
A:Accession: B91110
A:Molecule type: protein
A:Residues: 1-91 <IS2>
R:Baudier, J.; Gerard, D.
Biochemistry 22, 3360-3369, 1983
A:Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc C
A:Reference number: A90471; MUID:84000339
A:Contents: annotation; metal ion-binding properties
R:Marshak, D.R.; Umekawa, H.; Watterson, D.M.; Hidaka, H.
Arch. Biochem. Biophys. 240, 777-780, 1985
A:Title: Structural characterization of the calcium binding protein S100 from adipose ti
A:Reference number: A90075; MUID:85278169
A:Accession: A90075
A:Molecule type: protein
A:Residues: 1-91 <MAR>
A:Experimental source: adipose tissue
R:Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.
Biochem. J. 306, 551-555, 1995
A:Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia
A:Reference number: S54343; MUID:95194333
A:Accession: S54348
A:Molecule type: protein
A:Residues: 56-61, 'N', 63-79, 'V' <OKA>
C:Comment: The S-100 protein is composed of two related polypeptide chains, alpha and be
C:Comment: S-100 is also found in a variety of other tissues.
C:Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc ve
C:Comment: S-100 protein; calmodulin repeat homology <EF2>
C:Keywords: blocked amino end; brain; calcium binding; EF hand; zinc
F:48-80/Domain: calmodulin repeat homology <EF2>
F:1/Modified site: blocked amino end (Ser) (probably acetylated) #status experimental
```

```
F:18,21,23,26,31/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F:61,63,65,67,72/Binding site: calcium (Asp, Asp, Asp, Glu, Glu) #status predicted

alignment_scores:
Quality: 137.00 Length: 86
Ratio: 2.210 Gaps: 1
Percent Similarity: 72.093 Percent Identity: 33.721

alignment_block:
US-09-806-382a-1 x BCBOIB ..

Align seg 1/1 to: BCBOIB from: 1 to: 91

7 ACCGAGCTGAGAAAGCCTTGAACCTATCATGACGTCATACCAAGTA 56
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 SerGIuLeuGIuLysAlaValAlaLeuLLeAspValPheHIsGInTy 17
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 rSerGIyArggLuGIyAspLysHIsLysLeuLysSerGIuLeuLys 34
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 CTCCCTGATMAAGGGGAATTTCCATGCCGTCTACAGAGATGACCTGAGA 106
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 LuLeuLLeAsnSngLuLeuSerHIsPheLeuLuGIuLeuLysGIuIn 50
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 AAGGGTCAGAGCTGTGGTTCAAGAGTGGATATCAACAGCTGATGTCG 194
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 GIuValAlaAspLysValMetGIuThrLeuAspSerAspGIuAspGI 67
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 AGTTAACTTCAGAGATCTCTCTTCTGGTGATMAAGATGGCGGTGAG 244
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 uCysAspPheGInGIuPheMetAlaPheValAlaMetIleThrThraIac 84
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 CCGACAA 252
84 YSHISGIu 86

seq_name: pIrl:S35985

seq_documentation_block:
S-100 protein alpha chain - weatherfish
C:Species: Misgurnus fossilis (weatherfish)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S35985
R:Rivanekov, V.V.; Gerke, V.; Minhn, A.A.; Plessmann, U.; Weber, K.
Mech. Dev. 42, 151-158, 1993
A:Title: Transduction of Ca(2+) signals upon fertilization of eggs; Identification of
A:Reference number: S35985; MUID:94031845
A:Accession: S35985
A:Molecule type: protein
A:Residues: 1-95 <IVA>
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:49-81/Domain: calmodulin repeat homology <EF2>

alignment_scores:
Quality: 134.00 Length: 92
Ratio: 2.271 Gaps: 2
Percent Similarity: 64.130 Percent Identity: 34.783

alignment_block:
US-09-806-382a-1 x S35985 ..

Align seg 1/1 to: S35985 from: 1 to: 95

4 TTAGCCAGCTGGAGAAAGCCTTGAACCTATCATGACGTCATACCAAA 53
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 ValSerGIuLeuGIuSerAlaMetGIuSerLeuLLeLysValPheHIsH 17
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 GTACTCTCTGATMAAGGGGAATTTCCATGCCGTCTACAGAGATGACCTGA 103
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

17 rTyrSerSerLysGluGlyAspLysTyrLysLeuSerLysAlaGluLeuL 34  
104 AGAATTCCTAGAGACCGAGTGT..... 126  
34 ySerLeuLeuGlnGlyIuLeuAsnAspPheLeuSerAlaSerLysAsp 50  
127 CCTCAGTATATCAGAAAAGGGTGCAGACCTGTCTGTCAAGAGTTGGA 176  
51 PrometValValGluLysTlleMetSerAsp.....LeuAs 62  
177 TATCAACACTGATGGTGCAGTTAACTCCAGAGATTCTCATTTGTGTA 226  
62 pGluAsnGlnAspGlyGluValAspPheGlnGluPheValValLeuValA 79  
227 TAAAGATGGGCGTGCAGCCACACAAA 252  
79 laAlaLeuThrValAlaCysAsnGlu 87



OM of: US-09-806-382a-1 to: SwissProt\_40.\* out\_format : pfs

Date: Sep 9, 2002 3:22 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODL=frim+asp.model -DEV=xlp  
-O=/cgn2\_1/USPRO.spool/US09806382/rmat\_09092002.143849\_15565/app\_query.fasta.1.752  
-DB=SwissProt\_40 -OFMT=fastan -SUFFIX=rsf -GAPOP=12.000  
-GAPEXT=4.000 -MIMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -TGAOP=10.000 -TGAPEXT=0.500  
-DELOP=0.000 -DELEXT=7.000 -START=1 -MARGIN=blsnum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=Pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OFFPMT=Pfs  
-NOR=ext -HEAPSIZE=500 -MILEN=6 -MAXLEN=2000000000  
-USER=US09806382.@CGN1.1.62 -NCPU=6 -IOPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-806-382a-1

Query length: 282

Database: SwissProt\_40.\*

Database sequences: 105224

Database length: 38719550

Search time (sec): 32.310000

Score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
SwissProt_40:S108_HUMAN + 485.00	1000.28	5.5e-48	93	P05109 homo sapiens (human).	
SwissProt_40:S108_RAT + 320.00	659.13	5.8e-29	88	P50115 rattus norvegicus (rat).	
SwissProt_40:S108_MOUSE + 293.00	603.23	7.6e-26	88	P27005 mus musculus (mouse).	
SwissProt_40:S108_BOVIN + 170.00	355.03	1.1e-11	41	P28782 bos taurus (bovine).	
SwissProt_40:M126_CHICK + 169.50	344.94	1.4e-11	119	P28318 gallus gallus (chick).	
SwissProt_40:S102_BOVIN + 158.00	323.45	2.8e-10	91	P79105 bos taurus (bovine).	
SwissProt_40:S102_BOVIN + 156.50	317.86	4.3e-10	122	P28783 bos taurus (bovine).	
SwissProt_40:S102_HUMAN + 156.00	319.31	4.8e-10	91	P80511 homo sapiens (human).	
SwissProt_40:S102_HUMAN + 154.00	314.80	8.1e-10	95	P25815 homo sapiens (human).	
SwissProt_40:S102_PIG + 151.00	308.95	1.8e-09	91	P80310 sus scrofa (pig).	
SwissProt_40:S104_RAT + 148.00	302.56	4.0e-09	93	P35467 rattus norvegicus (rat).	
SwissProt_40:S104_BOVIN + 146.00	298.42	6.8e-09	93	P26630 bos taurus (bovine).	
SwissProt_40:S104_HUMAN + 144.00	294.28	1.2e-08	93	P23297 homo sapiens (human).	
SwissProt_40:S104_MOUSE + 139.00	283.93	4.4e-08	93	P26565 mus musculus (mouse).	
SwissProt_40:S108_BOVIN + 137.00	279.97	7.4e-08	91	P26368 bos taurus (bovine).	
SwissProt_40:S108_HUMAN + 137.00	279.97	7.4e-08	91	P42271 homo sapiens (human).	
SwissProt_40:S108_MOUSE + 137.00	279.97	7.4e-08	91	P50114 mus musculus (mouse).	
SwissProt_40:S108_RAT + 136.00	277.90	9.7e-08	91	P04631 rattus norvegicus (rat).	
SwissProt_40:S109_HUMAN + 133.50	270.82	1.9e-07	114	P06702 homo sapiens (human).	
SwissProt_40:S109_MOUSE + 131.50	266.83	3.2e-07	112	P31725 mus musculus (mouse).	
SwissProt_40:S111_MOUSE + 131.00	266.92	3.7e-07	98	P50143 mus musculus (mouse).	
SwissProt_40:S109_RAT + 130.50	264.76	4.2e-07	112	P50116 rattus norvegicus (rat).	
SwissProt_40:S107_BOVIN + 127.50	259.42	9.3e-07	101	P028050 bos taurus (bovine).	
SwissProt_40:S112_BOVIN + 127.00	250.25	1.0e-06	81	P077791 oryctolagus cuniculus.	
SwissProt_40:S104_HUMAN + 126.00	256.31	1.4e-06	101	P26447 homo sapiens (human).	
SwissProt_40:S104_BOVIN + 125.00	253.33	1.8e-06	100	P35466 bos taurus (bovine).	
SwissProt_40:S104_MOUSE + 124.00	252.17	2.4e-06	101	P07091 mus musculus (mouse).	
SwissProt_40:S105_MOUSE + 122.00	248.73	4.0e-06	93	P08894 mus musculus (mouse).	
SwissProt_40:S111_RABIT + 122.00	247.95	4.0e-06	102	P24480 oryctolagus cuniculus.	
SwissProt_40:S109_RABIT + 121.50	245.68	4.6e-06	118	P50117 oryctolagus cuniculus.	
SwissProt_40:S104_RAT + 121.00	245.36	5.2e-06	101	P05942 rattus norvegicus (rat).	
SwissProt_40:S111_CHICK + 121.00	245.96	5.2e-06	101	P24479 gallus gallus (chick).	
SwissProt_40:S105_HUMAN + 120.50	244.20	6.0e-06	110	P33763 homo sapiens (human).	
SwissProt_40:S102_BOVIN + 119.50	244.04	7.6e-06	78	P02633 bos taurus (bovine).	
SwissProt_40:S102_HUMAN + 118.50	241.13	1.0e-05	97	P026304 homo sapiens (human).	
SwissProt_40:S102_PIG + 115.50	236.76	2.2e-05	78	P026302 sus scrofa (pig).	
SwissProt_40:S106_RABIT + 113.00	230.37	4.3e-05	90	P30801 oryctolagus cuniculus.	
SwissProt_40:S101_HUMAN + 112.50	228.12	4.9e-05	92	P29377 homo sapiens (human).	
SwissProt_40:S101_ICPUP + 112.00	228.12	5.7e-05	78	P09106 ictalurus punctatus (cat).	
SwissProt_40:S102_BOVIN + 112.00	227.67	5.7e-05	97	P10463 bos taurus (bovine).	

seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S111_PIG + 112.00	227.50	5.7e-05	99	P31950 sus scrofa (pig).	cal	
SwissProt_40:S111_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00	200.92	0.0001	1407	P37709 oryctolagus cunic.		
seq_name: SwissProt_40:S108_HUMAN	seq_documentation_block:	ID	S108_HUMAN	STANDARD:	PRT:	93 AA.
SwissProt_40:S108_HUMAN + 111.00	224.93	7.5e-05	105	P31949 homo sapiens (human).		
SwissProt_40:S110_CHICK + 110.50	224.65	8.5e-05	96	P27003 gallus gallus (chick).		
SwissProt_40:S106_HUMAN + 110.00	224.16	9.6e-05	90	P06703 homo sapiens (human).		
SwissProt_40:TRH_RABIT + 110.00						

```

RT protein family, by MAD phasing at 1.9 A resolution.";
RL Acta Crystallogr. D 56:559-566(2000).
CC -1- FUNCTION: EXPRESSED BY MACROPHAGES IN CHRONIC INFLAMMATIONS. ALSO
CC EXPRESSED IN EPITHELIAL CELLS CONSTITUTIVELY OR INDUCED DURING
CC DERMATOSSES. MAY INTERACT WITH COMPONENTS OF THE INTERMEDIATE
CC FILAMENTS IN MONOCYTES AND EPITHELIAL CELLS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- MISCELLANEOUS: BINDS TWO CALCIUM IONS PER MOLECULE WITH AN
CC AFFINITY SIMILAR TO THAT OF THE S-100 PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y00378; CA68390.1; -
DR EMBL; X06234; CA29580.1; -
DR EMBL; M21005; AA36327.1; -
DR EMBL; A12027; CAA01001.1; -
DR PIR; A29764; BCHUCF.
DR PIR; A31848; A31848.
DR PIR; S13454; S13454.
DR PDB; 1MR8; 17-MAT-00.
DR Aarhus/Ghent-2DPAGE; 1003; IEF.
DR PMMA-2DPAGE; P05109; -
DR MIM; 123885; -
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_Cabp.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR Macrophage; Calcium-binding; 3D-structure.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 59 70 EF-HAND 2 (HIGH AFFINITY).
FT CONFLICT 79 93 GVAHKKSHESHKE -> AMOPTKAMKKATKSS
FT SEQUENCE 93 AA; 10834 MW; 78F589140B9CE166 CRC64;
SQ
alignment_scores:
Quality: 485.00 Length: 93
Ratio: 5.215 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-806-382a-1 x S108_HUMAN ..
Align seg 1/1 to: S108_HUMAN from: 1 to: 93
1 ATGTGACGAGCTGGAGAAAGCCCTGAACCTATCATCGACGCTACCA 50
|||||
1 MetLeuThGluLeuGluLysAlaLeuAsnSerIleIleAspValIYrHl 17
51 CAAGTACTCCCTGATTAAGGGGAATTCATCGCCGCTACAGGAGTACC 100
|||||
17 slYsYrSerLeuIleLysGlyAsnPhenIleAlaValIYrArGAspAspL 34
101 TGAAGAAATTCGTAGAGACCGAGTCCCTCAGTATATCAGGAAAGGGT 150
|||||
34 euLysYsLeuLeuGluThnGluCysProGlnIYrIleAqLysLysGly 50
151 GCAGAGCTGTGTTCAAGAGTTGATATCAACACTGATGTCACATTAA 200
|||||
51 AlaAspValIYrPhelYsGluLeuAspIleAsnThrAspGlyAlaValAs 67
202 CTTTCAGGAGTTCCTATTCTGGTGTGATTAAGATGGCGTGCGACCCACA 250
|||||

```

```

67 nPhenGluPhLeuLeuIleuValIleLysMetGlyValAlaAlaHsl 84
251 AAAAAAGCCATGAAGAAAGCCACCAAGAG 279
|||||
84 yslYsSerHisGluGluSerHisLysGlu 93
seq_name: SwissProt_40:S108_RAT
seq_documentation_block:
ID S108_RAT STANDARD; PRT; 88 AA.
AC P50115;
DT 01-OCT-1996 (Rel. 34, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calgranulin A (Migration inhibitory factor-related protein 8) (MRP-8)
DE (P8).
GN S100A8 OR MRP8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEWIS/N; TISSUE=peritoneal cavity;
RX MEDLINE=93343942; PubMed=8343166;
RA Imanishi T., Uchida I., Wahl S.M., McCartney-Francis N.;
RT "Expression and cloning of migration inhibitory factor-related
RT protein (MRP) 8 and MRP14 in arthritis-susceptible rats.";
RL Biochem. Biophys. Res. Commun. 194:819-825(1993).
RN [2]
RP SEQUENCE, AND ACETYLATION.
RC TISSUE=Spleen;
RX MEDLINE=98249881; PubMed=9570842;
RA Raftery M.J., Geecy C.L.;
RT "Identification of posttranslational modifications and cDNA sequencing
RT errors in the rat S100 proteins MRP8 and 14 using electrospray
RT ionization mass spectrometry.";
RL Anal. Biochem. 258:285-292(1998).
CC -1- FUNCTION: HAS MAXIMAL CHEMOTACTIC ACTIVITY FOR NEUTROPHILS. MAY
CC PLAY AN IMPORTANT ROLE IN CELLULAR IMMUNE REACTIONS.
CC -1- SUBUNIT: HOMODIMER AND/OR HETERODIMER.
CC -1- MASS SPECTROMETRY: MW=10149; MW_ERR=2; METHOD=Electrospray.
CC -1- MISCELLANEOUS: BINDS TWO CALCIUM IONS PER MOLECULE WITH AN
CC AFFINITY SIMILAR TO THAT OF THE S-100 PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L18891; AAA41637.1; -
DR HSSP; P04631; 1BAC.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_Cabp.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR SMART; SM00054; EFh; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR Calcium-binding; Chemotaxis; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1
FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 58 69 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT CONFLICT 72 72 V -> A (IN REF. 1).
SQ SEQUENCE 88 AA; 10107 MW; 235CE67BA9602D1E CRC64;

```



```

alignment_scores:
  Quality: 320.00      Length: 87
  Ratio: 3.951        Gaps: 0
  Percent Similarity: 93.103      Percent Identity: 63.218

alignment_block:
US-09-806-382a-1 x S108_RAT ..

Align seg 1/1 to: S108_RAT from: 1 to: 88

7 ACCGAGCTGGAGAAAGCTTGAACCTATATCACTGACGAGTACACAGATA 56
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 ThrGluLeuGluLysAlaLeuSerSrnValIleGluValTyrHisAsnTyr 18
57 CTCCTGATTAAGGGAATTTCCATCCGCTGACAGGATGACATGAGA 106
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
18 rSerGlyLeuLysGlyAsnHisAlaLeuTyrTrpGlyAspPheArgL 35
107 AATTGCTAGAGACGAGTGTCTCTCAGTATATCAGAGAAAGGCTGACAG 156
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
35 yMeValThrThrGluLysProGlnPheValGlnAsnLysAsnThrGlu 51
157 GTCGTGTCGAAAGAGTGTGATCAACACTGATGTCGAGTAACTTCCA 206
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 SerLeuPheLysGluLeuAspValAlaSerAspAsnAlaIleAsnPhcGl 68
207 GGAGTTCCTCATCTGCTGATTAAGATGGCGTGGCAGCCCAAAAAA 256
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
68 uGluPheLeuValLeuValIleArgValGlyValAlaAlaHisLysAspS 85
257 GCCATGAAGAA 267
|||||:|||||
85 eRHISLysGlu 88

seq_name: SwissProt_40:S108_MOUSE

seq_documentation_block:
ID S108_MOUSE STANDARD; PRT; 88 AA.
AC P27005; P31724;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calgranulin A (Migration inhibitory factor-related protein 8) (MRP-8)
DE (P8) (Leukocyte L1 complex light chain) (Chemotactic cytokine CP-10)
DE (PRO-inflammatory S100 cytokine).
GN S100A8 OR CAGA OR MRP8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=92223366; PubMed=1373330;
RA Lagasse E., Weissman I.L.;
RT "Mouse MRP8 and MRP14, two intracellular calcium-binding proteins
RT associated with the development of the myeloid lineage.";
RL Blood 79:1907-1915(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93203618; PubMed=8454868;
RA Lackmann M., Rajasekariah P., Iismaa S.E., Jones G., Cornish C.J.,
RA Hu S., Simpson R.J., Moritz R.L., Geczy C.L.;
RT "Identification of a chemotactic domain of the pro-inflammatory S100
RT protein CP-10.";
RL J. Immunol. 150:2981-2991(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; TISSUE=Liver;
RX MEDLINE=96190772; PubMed=8611640;
RA Nacken W.K.F., Manitz M.P., Sorg C.;
RT "Molecular characterisation of the genomic locus of the mouse MRP8
RT gene.";
```

```

RL Biochim. Biophys. Acta 1315:1-5(1996).
RN [4]
RP SEQUENCE OF 1-76.
RC TISSUE=Spleen;
RA MEDLINE=92218405; PubMed=1559987;
RX Lackmann M., Cornish C.J., Simpson R.J., Moritz R.L., Geczy C.L.;
RT "Purification and structural analysis of a murine chemotactic
RT cytokine (CP-10) with sequence homology to S100 proteins.";
RL J. Biol. Chem. 267:7499-7504(1992).
CC -1- FUNCTION: HAS MAXIMAL CHEMOTACTIC ACTIVITY FOR NEUTROPHILS. MAY
CC PLAY AN IMPORTANT ROLE IN CELLULAR IMMUNE REACTIONS.
CC -1- SUBUNIT: HOMODIMER AND/OR HETERODIMER.
CC -1- MISCELLANEOUS: BINDS TWO CALCIUM IONS PER MOLECULE WITH AN
CC AFFINITY SIMILAR TO THAT OF THE S-100 PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M83218; AAB07229.1; -.
DR EMBL; S57123; AAB25840.1; -.
DR EMBL; X87966; CAA61204.1; -.
DR PIR; A42488; A42488.
DR HSSP; P04631; 1BAC.
DR MGD; MGI:86244; S100a8.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001751; S100_Cabp.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Calcium-binding; Chemotaxis.
FT INIT_MET 0 0
FT CA_BIND 19 32 EF_HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 58 66 EF_HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT CONFLICT 56 56 E -> D (IN REF. 3).
SQ SEQUENCE 88 AA; 10163 MW; 9B32BDBE53CFE728 CRC64;

alignment_scores:
  Quality: 293.00      Length: 87
  Ratio: 3.756        Gaps: 0
  Percent Similarity: 89.655      Percent Identity: 58.621

alignment_block:
US-09-806-382a-1 x S108_MOUSE ..

Align seg 1/1 to: S108_MOUSE from: 1 to: 88

7 ACCGAGCTGGAGAAAGCTTGAACCTATATCACTGACGAGTACACAGATA 56
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2 rSerAnIleGlnGlyAsnHisAlaLeuTyrTrpGlyAspPheArgL 35
18 rSerAnIleGlnGlyAsnHisAlaLeuTyrTrpGlyAspPheArgL 35
107 AATTGCTAGAGACGAGTGTCTCTCAGTATATCAGAGAAAGGCTGACAG 156
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
35 yMeValThrThrGluLysProGlnPheValGlnAsnLysAsnThrGlu 51
157 GTCGTGTCGAAAGAGTGTGATCAACACTGATGTCGAGTAACTTCCA 206
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 AsnLeuPheArgLysGluLeuAspPheValAlaIleAsnPhcGl 68
207 GGAGTTCCTCATCTGCTGATTAAGATGGCGTGGCAGCCCAAAAAA 256
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

68 uc1upheleuAlaMetValIleuysValGlyValAlaSerHisLysAsps 85  
 257 GCCATGAGAA 267  
 |||:|||||  
 85 erHisLysGlu 88

seq\_name: SwissProt\_40:S108\_BOVIN

seq\_documentation\_block:  
 ID S108\_BOVIN STANDARD: PRT: 41 AA.  
 AC P28782:  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Calgranulin A (Neutrophil cytosolic 7 kDa protein) (P7) (BEE11)  
 DE (Fragment).  
 GN S100A8.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Neutrophils;  
 RX MEDLINE=92304974; PubMed=1610833;  
 RA Dianoux A.-C., Stasia M.-J., Gagnon J., Vignais P.V.;  
 RT "The 23-kilodalton protein, a substrate of protein kinase C, in  
 RT bovine neutrophil cytosol is a member of the S100 family.";  
 RL Biochemistry 31:5898-5905(1992).  
 RN [2]  
 RP SEQUENCE OF 1-29.  
 RC TISSUE=Esophageal epithelium;  
 RX MEDLINE=93280230; PubMed=8505356;  
 RA Tang T.K., Hong T.-M., Lin C.-Y., Lai M.-L., Liu C.H.L., Lo H.-J.,  
 RA Wang M.-E., Chen L.B., Chen W.-T., Ip W., Lin D.C., Lin J.-C.,  
 RA Lin S., Sun T.-T., Wang E., Wang J.L., Wu R., Wu C.-W., Chien S.;  
 RT "Nuclear proteins of the bovine esophageal epithelium. I. Monoclonal  
 RT antibody W2 specifically reacts with condensed nuclei of  
 RT differentiated superficial cells.";  
 RL J. Cell Sci. 104:237-247(1993).  
 CC -1- SUBUNIT: DISULFIDE LINKED HETERODIMER OF A 7/11 kDa AND A 22/23  
 CC kDa SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; LOOSELY ASSOCIATED TO THE  
 CC CYTOSKELETON.  
 CC -1- TISSUE SPECIFICITY: FOUND ESSENTIALLY IN PHAGOCYtic CELLS.  
 CC -1- MISCELLANEOUS: BINDS TWO CALCIUM IONS PER MOLECULE WITH AN  
 CC AFFINITY SIMILAR TO THAT OF THE S-100 PROTEINS (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.  
 DR PIR: B42628; B42628.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR001751; S100\_CaBP.  
 DR Pfam: PF01023; S100; 1.  
 DR PROSITE: PS00018; EF\_HAND; PARTIAL.  
 DR PROSITE: PS00303; S100\_CaBP; PARTIAL.  
 KW Calcium-binding.  
 FT CA\_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).  
 FT CONFLICT 10 11 DS -> NI (IN REF. 2).  
 FT NON\_TER 41 41  
 SQ SEQUENCE 41 AA; 4770 MW; 21C78BBF3ABDAB55 CRC64;

alignment\_scores:  
 Quality: 170.00 Length: 41  
 Ratio: 4.595 Gaps: 0  
 Percent Similarity: 90.244 Percent Identity: 78.049

alignment\_block:  
 US-09-806-382A-1 x S108\_BOVIN ..

Align seg 1/1 to: S108\_BOVIN from: 1 to: 41

1 ATGTGACCGAGCTGGAGAAAGCCTTGACCTATCATCGACGCTACCA 50  
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||  
 1 MetLeuThrAspLeuGlu\*\*AlaIleAspSerLeuIleAspValTyrH 17  
 51 CAAGTACTCCCTGATAAAGGGGAATTCATGCGCTACAGGAGTACGCC 100  
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||  
 17 sLySTyrSerLeu\*\*\*LysGlyAsnTyrHisAlaValTyr\*\*AspAspL 34  
 101 TGAAGAAATTCCTAGAGACCGAG 123  
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||  
 34 euLys\*\*\*LeuLeuGluThrGlu 41

seq\_name: SwissProt\_40:M126\_CHICK

seq\_documentation\_block:  
 ID M126\_CHICK STANDARD: PRT: 119 AA.  
 AC P28318;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Protein MRP-126.  
 DE Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WHITE LEGHORN; TISSUE=Bone marrow;  
 RX MEDLINE=92195690; PubMed=1549365;  
 RA Nakano T., Graf T.;  
 RT "Identification of genes differentially expressed in two types of  
 RT v-myb-transformed avian myelomonocytic cells.";  
 RL Oncogene 7:527-534(1992).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN V-MYB-TRANSFORMED MYELOMONOCYTIC  
 CC CELLS.  
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X61200; -. NOT\_ANNOTATED\_CDS.  
 DR HSSP: P04271; 10WO.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR001751; S100\_CaBP.  
 DR Pfam: PF00036; ehand; 1.  
 DR Pfam: PF01023; S100; 1.  
 DR PROSITE: PS00018; EF\_HAND; 1.  
 DR PROSITE: PS00303; S100\_CaBP; 1.  
 KW Calcium-binding.  
 FT CA\_BIND 29 42 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).  
 FT CA\_BIND 72 83 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).  
 FT CONFLICT 119 14065 MW; 2D268DAF6309AD7A CRC64;  
 SQ SEQUENCE 119 AA; 14065 MW; 2D268DAF6309AD7A CRC64;

alignment\_scores:  
 Quality: 169.50 Length: 100  
 Ratio: 2.457 Gaps: 2  
 Percent Similarity: 69.000 Percent Identity: 36.000

alignment\_block:  
 US-09-806-382A-1 x M126\_CHICK ..

Align seg 1/1 to: M126\_CHICK from: 1 to: 119

4 TTGACCGAGCTGGAGAAAGCCTTGACCTATCATCGACGCTACCA 53  
 |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

```

11 LeuSerGIuLeuGIuLSaIaIleAspValIleIleAspValPheHisGI 27
54 GTATCCCTCGTAAAGGGAATTTCATGCCGCTACAGGAGTACCTGA 103
||||| :||| :||| :|||
27 nTYSerArgArgIGluGLYAspLysAspThrLeuThrArgLysGIuLeu 44
104 AGAAATTCAGAGACCGAGTGTCTCAGTATTCAGG..... 141
|| :||| :||| :|||
44 ySLeuLeuIleGIuLysGIuLeuAlaSnTyrLeuLysValLysAsn 60
142 AAAAAGGTCGACAGCGTGTCTCAAGAGTGTATATCACTACAGTACG 191
||| :||| :||| :|||
61 GluValSerIleAspGlnIlePheLysAspLeuAspAsnLysAspGI 77
192 TGCAGTTAACTCCAGAGTTCCTCATTCGTGTATTAAGATGGCGTGG 241
||| :||| :||| :|||
77 nGIuLeuSerPheGIuLysValMetLeuLeuIleIleArgValThrVal 94
242 CAGCCCAACAA..... AAAAGCATGACAAACCCACAA 276
||| :||| :||| :|||
94 IaThrHisGIuHisLeuHisPheCysGIuAspHisGIuHisGIu 110
seq_name: SwissProt_40:S112_BOVIN

seq_documentation_block:
ID S112_BOVIN STANDARD; PRT; 91 AA.
AC P79105;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Calgranulin C (CAGC) (Calcium-binding protein in amniotic fluid 1)
DE (CAAF1) (RAGE binding protein).
OS S100A12 or CAAF1.
GN Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oesophagus;
RX MEDLINE=96298783; PubMed=8718672;
RA Hitomi J., Yamaguchi K., Kikuchi Y., Kimura T., Maruyama K.,
RA Nagasaki K.;
RT "A novel calcium-binding protein in amniotic fluid, CAAF1: its
RT molecular cloning and tissue distribution.";
RL J. Cell Sci. 109:805-815(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=9325504; PubMed=10399917;
RA Hofmann M.A., Drury S., Fu C., Qu W., Taguchi A., Lu Y., Avila C.,
RA Kamhan N., Biehn A., Navroth P., Neurath M.F., Slattery T.,
RA Beach D., McClary J., Nagashima M., Morser J., Stern D.,
RA Schmidt A.M.;
RT "RAGE mediates a novel proinflammatory axis: a central cell surface
RT receptor for S100/calgranulin polypeptides.";
RL Cell 97:889-901(1999).
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
CC EMBL: D49548; BAA08496.1; -
CC DR EMBL: AF011757; AAB65423.1; -
CC DR HSSP: P02633; IBOC.
CC DR InterPro: IPR002048; EF-hand.

```

```

DR InterPro: IPR001751; S100_CaBP.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF01023; S_100; 1.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00303; S100_CaBP; 1.
KW Calcium-binding; zinc; Metal-binding.
FT INT_MET 0 0 BY SIMILARITY.
FT CA_BIND 18 31 EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).
FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (BY SIMILARITY).
SQ SEQUENCE 91 AA; 10554 MW; 66FBC3CB0354482 CRC64;

alignment_scores:
Quality: 158.00 Length: 91
Ratio: 2.508 Gaps: 1
Percent Similarity: 69.231 Percent Identity: 36.264

alignment_block:
US-09-806-382a-1 x S112_BOVIN ..
Align seg 1/1 to: S112_BOVIN from: 1 to: 91

7 AGCGAGCTGGAGAAAGCCTTGACCTATTCATCGACGCTACCCACAGTA 56
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1 ThrLysLeuGlnAspHisLeuGIuLysValIleLeuAsnIlePheHisGIu 17
57 CTCCTGTATTAAGGGAATTTCATCCGCTACAGGATGACCTGAGA 106
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
17 rSerValArgValAlGIuHisPheAspThrLeuAsnLysArgGIuLeuLysG 34
107 AATGCTAGACACCGAGTGTCTCAGTATTCAGG.....AAA 144
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
34 InLeuIleThrLysGIuLeuProLysThrLeuGlnAsnThrLysAspGIu 50
145 AAGGTCGACAGCTGTGGTCAAGAGTTCATATCAACACTGATGATGCG 194
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
51 ProThrIleAspLysIlePheGlnAspLeuAspAlaAspLysAspGIu 67
195 AGTTAACTTCAGAGATTCCTCATTCGTGTATTAAGATGGCGCTGCGAG 244
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
67 aValSerPheGIuLysPheValIleValSerArgValLeuLysThrA 84
245 CCCACAAAAAAGCCATGACAA 267
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
84 IaHisIleAspIleHisGIu 91

seq_name: SwissProt_40:S109_BOVIN

seq_documentation_block:
ID S109_BOVIN STANDARD; PRT; 122 AA.
AC P28783;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calgranulin B (Neutrophil cytosolic 23 kDa protein) (P23) (BEE22)
DE (Fragment).
GN S100A9.
OC Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Oesophageal epithelium;
RX MEDLINE=93280230; PubMed=8505358;
RA Tang T.K., Hong T.-M., Lin C.-Y., Lai M.-L., Liu C.H.L., Lo H.-J.,
RA Wang M.-E., Chen L.B., Chen W.-T., Ip W., Lin D.C., Lin J.-C.,
RA Lin S., Sun T.-T., Wang E., Wang J.-L., Wu R., Wu C.-W., Chien S.;
RT "Nuclear proteins of the bovine esophageal epithelium. I. Monoclonal
RT antibody W2 specifically reacts with condensed nuclei of
RT differentiated superficial cells.";

```

```
seq_documentation_block:
  ID      S112_HUMAN      STANDARD;          PRT;          91 AA
  AC      P80511;
  DT      01-OCT-1996 (Rel. 34, Created)
  DT      01-OCT-1996 (Rel. 34, Last sequence update)
```

DE 01-MAR-2002 (Rel. 41, last annotation update)  
DE Calgranulin C (CAGC) (p6) (CGRP) (Neutrophil S100 protein) (Calcium-  
DE binding protein in amniotic fluid 1) (CAAF1) (p6).  
GN S100A12.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97138564; PubMed=8985590;  
RA Wicki R., Marenholz I., Mische D., Schaefer B.W., Heizmann C.W.;  
RT "Characterization of the human S100A12 (calgranulin C, p6, CAAF1,  
RT CGRP) gene, a new member of the S100 gene cluster on chromosome  
RL 1q21".  
RL Cell Calcium 20:459-464(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96192053; PubMed=8619860;  
RA Yamamura T., Hitomi J., Nagasaki K., Suzuki M., Takahashi E.,  
RA Saito S., Tsubata T., Yamaguchi K.;  
RT "Human CAAF1 gene -- molecular cloning, gene structure, and chromosome  
RT mapping".  
RL Biochem. Biophys. Res. Commun. 221:356-360(1996).  
RN [3]  
RP SEQUENCE.  
RX MEDLINE=96192069; PubMed=8619876;  
RA Marti T., Ertmann K.D., Gallin M.Y.;  
RT "Host-parasite interaction in human oncocarcinosis: identification  
RT and sequence analysis of a novel human calgranulin C".  
RL Biochem. Biophys. Res. Commun. 221:454-458(1996).  
RN [4]  
RP SEQUENCE.  
RX TISSUE=Neutrophils;  
RC MEDLINE=96332419; PubMed=8769108;  
RA Ilg E.C., Troxler H., Buergisser D.M., Kuster T., Markert M.,  
RA Guignard F., Hunziker P., Birchler N., Heizmann C.W.;  
RT "Amino acid sequence determination of human S100A12 (p6, calgranulin  
RT C, CGRP, CAAF1) by tandem mass spectrometry".  
RL Biochem. Biophys. Res. Commun. 225:146-150(1996).  
RN [5]  
RP SEQUENCE OF 1-20.  
RX MEDLINE=95351965; PubMed=7626002;  
RA Guignard F., Maue J., Markert M.;  
RT "Identification and characterization of a novel human neutrophil  
RT protein related to the S100 family".  
RL Biochem. J. 309:395-401(1995).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).  
RX MEDLINE=21065388; PubMed=11134923;  
RA Moroz O.V., Antson A.A., Murshudov G.N., Matland N.J., Dodson G.G.,  
RA Wilson K.S., Skidshoj I., Lukandov E.M., Bronstein I.B.;  
RT "The three-dimensional structure of human S100A12.";  
RL Acta Crystallogr. D 57:20-29(2001).  
RN [7]  
RP SUBUNIT: HOMODIMER.  
CC -1- TISSUE SPECIFICITY: MONOCYTES AND LYMPHOCYTES.  
CC -1- MASS SPECTROMETRY: MW=10444; METHOD-Electrospray.  
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X97859; CAA66453.1; -  
DR EMBL; X98288; CAA66934.1; -  
DR EMBL; X98289; CAA66934.1; JOINED.  
DR EMBL; X98290; CAA66934.1; JOINED.  
DR EMBL; X98299; CAB94792.1; -

DR EMBL: X98290; CAB94792.1; JOINED.  
 DR EMBL: D49549; BAA08497.1; -  
 DR EMBL: D83664; BAA12036.1; -  
 DR EMBL: D83657; BAA12030.1; -  
 DR PDB: 1E8A; 08-JAN-01.  
 DR MIM: 603112; -  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR001751; S100\_CABP.  
 DR Pfam: PF00036; ehand; 1.  
 DR Pfam: PF01023; S\_100; 1.  
 DR PROSITE: PS00018; EF\_HAND; FALSE\_NEG.  
 DR PROSITE: PS00303; S100\_CABP; 1.  
 DR Calcium-binding; Zinc; Metal-binding; 3D-structure.  
 FT INIT\_MEF 0  
 FT CA\_BIND 18 31 EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).  
 FT CA\_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (BY  
 FT SEQUENCE 91 AA; 10444 MW; 325685EA8695F6B7 CRC64;  
 SIMILARITY).

## alignment\_scores:

Quality: 156.00 Length: 95  
 Ratio: 2.516 Gaps: 2  
 Percent Similarity: 65.263 Percent Identity: 38.947

## alignment\_block:

US-09-806-382a-1 x S112\_HUMAN ..

Align seg 1/1 to: S112\_HUMAN from: 1 to: 91

```

7  ACCGACCTGGAGAAAGCCTGGAACCTATCATCGACGCTTACCACAACTA 56
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
1  ThrLysLeuGluGluHisLeuGluGlyIleValAlaSnIlePheHisGlnTy 17
57  CTCCTCATATAAGGGATATTCATGCGCTTACAGGAGTACGACCTGACAGA 106
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
17  rSerValArgLysGlyHisPheAspThrLeuSerLysGlyGluLeuLysG 34
107  AATTGCTAGACACCGAGTCTCCATCATATATGAG.....AAA 144
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
34  ILeuLeuThrLysGlyLeuAlaIAsnThrIleLysAsnIleLysAspLys 50
145  AAGGCTGCAGACGCTGTGTTCAAGAGTGGATATCAACACATGATGGTCT 194
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
51  AlAlaValIleAspGluIlePheGlnGlyLeuAspAlaAsnGlnAspGluG 67
195  AGTTAACTTCACAGAGTTCCTCATTCGTGGTATTAAGATGGCGCTGGCAG 244
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
67  nValAspPheGlnGluPheIleSerLeuVal.....AlaIleA 80
245  CCCACAAAAAAGCCATGAAGAAAGCCACAAAGAG 279
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
80  lAlaLeuLysAlaAlaIAsnTyThrIAsnThrHisLysGlu 91

```

seq\_name: SWISSPROT\_40:S10E\_HUMAN

## seq\_documentation\_block:

ID S10E\_HUMAN STANDARD; PRT; 95 AA.  
 AC P25815;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE S100P protein.  
 GN S100P OR S100E.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=92339442; PubMed=1633809;  
 RA Becker T., Gerke V., Kube E., Weber K.;

RT "S100P, a novel Ca(2+)-binding protein from human placenta. cDNA  
 RT cloning, recombinant protein expression and Ca2+ binding  
 RT properties.";  
 RL Eur. J. Biochem. 207:541-547(1992).  
 RN [2]

RP SEQUENCE OF 1-91.  
 RC TISSUE=Placenta;  
 RX MEDLINE=92171935; PubMed=1540168;  
 RA Emoto Y., Kobayashi R., Akatsuka H., Hidaka H.;  
 RT "Purification and characterization of a new member of the S-100  
 RT protein family from human placenta.";  
 RL Biochem. Biophys. Res. Commun. 182:1246-1253(1992).  
 CC -1- MISCELLANEOUS: THIS PROTEIN BINDS TWO CALCIUM IONS.  
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed, usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL: X65614; CAA46566.1; -
DR PIR: S24146; S24146.
DR HSSP: P02638; 1CFP.
DR MIM: 600614; -
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_CABP.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF01023; S_100; 1.
DR PROSITE: PS00018; EF_HAND; FALSE_NEG.
DR PROSITE: PS00303; S100_CABP; 1.
KW Calcium-binding; Placenta.
FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY).
FT CONFLICT 32 32 E -> T (IN REF. 2).
FT CONFLICT 44 44 F -> E (IN REF. 2).
FT SEQUENCE 95 AA; 10400 MW; 786BE3FE3FAC6C61 CRC64;

```

## alignment\_scores:

Quality: 154.00 Length: 87  
 Ratio: 2.525 Gaps: 1  
 Percent Similarity: 70.115 Percent Identity: 37.931

## alignment\_block:

US-09-806-382a-1 x S10E\_HUMAN ..

Align seg 1/1 to: S10E\_HUMAN from: 1 to: 95

```

4  TTGACCGAGTGGAGAAAGCCTGGAACCTATCATCGACGCTTACCACAA 53
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
1  MethrGluLeuGluThrAlaMetGlyMetIleLeuAspValIlePheSerAr 17
54  GTACTCCCTGATTAAGGGGAATTCATGCGCTTACAGGAGTACGACCTGA 103
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
17  gTyrSerLysSerLysGlySerThrGlnThrLeuThrLysGlyGluLeuL 34
104  AGAATGCTGACAGACGAGTGCCTCAGTATATACAG..... 141
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
34  yValIleuMetGlnLysGluLeuProGlyPheLeuGlnSerGlyLysASP 50
142  AAAAAGGCTGCAGACGCTGTGTTCAAGAGTGGATATCAACACTGATGG 191
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
51  LysAspAlaValAspLysLeuLeuLysAspLeuAspAlaAsnGlnAspAl 67
192  TGCAGTTAACTTCACAGAGTTCCTCATTCGTGGTATTAAGATGGCGCTGG 241
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
67  aGlnValAspPheSerLysGluPheIleValAlaAlaIleThrSerA 84
242  CAGCCACAA 252

```

|||||  
84 lacysHs1sLys 87

seq\_name: SwissProt\_40:S112\_PIG

seq\_documentation\_block:  
ID S112\_PIG STANDARD: PRT: 91 AA.

AC P80310:

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Calgranulin C (CAGC).

GN S100A12.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

OC NCBI\_Taxid=9623;

OX [1]

RN

RP SEQUENCE.

RC TISSUE=Granulocyte;

RX MEDLINE=95050708; PubMed=7961855;

RA Dell'Angelica E.C., Schleicher C.H., Santome J.A.;

RT "Primary structure and binding properties of calgranulin C, a novel

RT S100-like calcium-binding protein from pig granulocytes.";

RL J. Biol. Chem. 269:28929-28936(1994).

CC -1- TISSUE SPECIFICITY: FOUND ESSENTIALLY IN GRANULOCYTES WITH SMALL

CC AMOUNTS FOUND IN LYMPHOCYTES.

CC -1- MISCELLANEOUS: IN THE ABSENCE OF ZINC BINDS ONE CALCIUM ION PER

CC MOLECULE. IN THE PRESENCE OF ZINC BINDS TWO CALCIUM IONS PER

CC

CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.

CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.

DR HSSP: P02632; ICBL

DR InterPro: IPR002048; EF-hand.

DR InterPro: IPR001751; S100\_CaBP.

DR Pfam: PF00036; ehand; 1.

DR PROSITE: PS00018; EF\_HAND; FALSE\_NEG.

DR PROSITE: PS00303; S100\_CaBP; 1.

KM Calcium-binding; zinc; Metal-binding.

FT CA\_BIND 18 31 EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).

FT CA\_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (BY

FT CA\_BIND 10614 MW; B4204461A32D7FCE CRC64;

SO SEQUENCE 91 AA; 10614 MW; B4204461A32D7FCE CRC64;

Alignment\_scores:

Quality: 151.00 Length: 91

Ratio: 2.323 Gaps: 2

Percent Similarity: 71.429 Percent Identity: 35.165

alignment\_block:

US-09-806-382a-1 x S112\_PIG ..

Align seg 1/1 to: S112\_PIG from: 1 to: 91

```

7 ACCGAGCTGGAGAAGCCTTGAAGCTGATGAGAGCTACCAAGTA 56
|||||
1 ThrLysLeuGluAspHisLeuGluGlyLeuLeuAsnIlePheHisGln 17
57 CTCCTGATTAAGGAGATTTCATGCCCTCTACAGAGGATGACCTGAGA 106
|||||
17 rSerValArgLeuGlyHisTyrAspThrLeuLeuLysArgGluLeuLys 34
107 AATTGCTAGAGACCGAGTCTCTCACTATATACAGAAA.....AAG 147
|||||
34 ILeuIleThrLysGluLeuProAsnThrLeuLysAsnThrLysAspGln 50
148 GGTGCA...GACGTCGTCAAGAGTGAATATCAACACAGATGCTGC 194
|||||
51 GlyThrIleAspLysIlePheGlnAsnLeuAspAlaAsnGlnAspGlu 67
195 AGTTAACTTCAGAGAGTTCTCATTTCTGGATTAAGATGGCGCTGAC 244

```

|||||  
245 CCCACAAAAAAGCCATGAGAA 267  
67 nValSerPheLysGluPheValValLeuValThrAspValLeuIleThr 84  
84 IaHisAspAsnIleHisLysGln 91

seq\_name: SwissProt\_40:S10A\_RAT

seq\_documentation\_block:  
ID S10A\_RAT STANDARD: PRT: 93 AA.

AC P35467;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE S-100 protein, alpha chain.

GN S100A1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC NCBI\_Taxid=10116;

OX [1]

RN

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RA Song W.;

RL submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 10-93 FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=92076235; PubMed=1742602;

RA Zimmer D.B., Song W., Zimmer W.E.;

RT "Isolation of a rat S100 alpha cDNA and distribution of its mRNA in

RT rat tissues.";

RL Brain Res. Bull. 27:157-162(1991).

CC -1- FUNCTION: WEAKLY BINDS CALCIUM BUT BINDS ZINC VERY TIGHTLY-

CC DISTINCT BINDING SITES WITH DIFFERENT AFFINITIES EXIST FOR BOTH

CC IONS ON EACH MONOMER. PHYSIOLOGICAL CONCENTRATIONS OF POTASSIUM

CC ION ANTAGONIZE THE BINDING OF BOTH DIVALENT CATIONS, ESPECIALLY

CC AFFECTING HIGH-AFFINITY CALCIUM-BINDING SITES.

CC -1- SUBUNIT: DIMER OF EITHER TWO ALPHA CHAINS, OR TWO BETA CHAINS, OR

CC ONE ALPHA AND ONE BETA CHAIN.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- TISSUE SPECIFICITY: ALTHOUGH PREDOMINANT AMONG THE WATER-SOLUBLE

CC BRAIN PROTEINS, S-100 IS ALSO FOUND IN A VARIETY OF OTHER TISSUES.

CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.

CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-slb.ch/announce/>

CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).

CC -----

DR EMBL: U26358; AAB53657.1; -

DR EMBL: U26357; AAB53657.1; JOINED.

DR EMBL: S68809; AAB20539.2; -

DR HSSP: P04631; 1B4C.

DR InterPro: IPR002048; EF-hand.

DR InterPro: IPR001751; S100\_CaBP.

DR Pfam: PF00036; ehand; 1.

DR PROSITE: PS001023; S\_100; 1.

DR PROSITE: PS00303; S100\_CaBP; 1.

KM Calcium-binding; zinc; Metal-binding.

FT CA\_BIND 0 0 EF-HAND 1 (LOW AFFINITY).

FT CA\_BIND 19 32 EF-HAND 2 (HIGH AFFINITY).

FT CA\_BIND 62 73 N -> H (IN REF. 2).

FT CONFLICT 13 13 K -> R (IN REF. 2).

FT CONFLICT 56 56

SO SEQUENCE 93 AA; 10429 MW; D2959A95EC0651A9 CRC64;

```

alignment_scores:
  Quality: 148.00      Length: 92
  Ratio: 2.277        Gaps: 2
  Percent Similarity: 70.652      Percent Identity: 39.130

alignment_block:
US-09-806-382a-1 x S10A_RAT ..

Align seg 1/1 to: S10A_RAT from: 1 to: 93

7 ACCGAGCTGGAGAAAGCCTTGAACCTATCATCGACGTCTACCAAGTA 56
  ::::::::::::::::::::|::::::::::::::::::|::
2 SerGluLeuGlnThrAlaMetGlnThrLeuIleAsnValPheHisAlaH1 18
  ::::::::::::::::::::|::::::::::::::::::|::
57 CTCCTCATTAAGGAGGAAATTTCCATGCCGCTACAGGATGACGTGAAGA 106
  ::::::::::::::::::::|::::::::::::::::::|::
18 sSerGlyLysGlnGlyAspLysTyrLysLeuSerLysLysGlnLeuLysA 35
  ::::::::::::::::::::|::::::::::::::::::|::
107 AATTGCTAGAGACCGAGTGTCTCTCAGTATATCAGG...AAAAGGTTGCA 153
  ::::::::::::::::::::|::::::::::::::::::|::
35 sPheLeuGlnThrGlnLeuSerPheLeuAspValGlnLysAspAla 51
  ::::::::::::::::::::|::::::::::::::::::|::
154 GACGTC.....TGTTCAAGAGTTGATATCAACACTGATGAGTGC 194
  ::::::::::::::::::::|::::::::::::::::::|::
52 AspAlaValAspLysIleMetLysGlnLeuAspGlnAsnGlyAspGlyG1 68
  ::::::::::::::::::::|::::::::::::::::::|::
195 AGTTAACTTCCAGAGAGTTCTCTGATTAAGATGGCGGTGCAG 244
  ::::::::::::::::::::|::::::::::::::::::|::
68 uValAspPheGlnGlnPheValIleValIleValAlaIleuThrValAlaC 85
  ::::::::::::::::::::|::::::::::::::::::|::
245 CCCACAAAAAAGCCATGAAGAAAGC 270
  ::::::::::::::::::::|::::::::::::::::::|::
85 ysAsnAsnPhePheTrpGlnAsnSer 93
  ::::::::::::::::::::|::::::::::::::::::|::

seq_name: SwissProt_40:S10A_BOVIN
seq_documentation_block:
ID S10A_BOVIN STANDARD; PRT; 93 AA.
AC P02639;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-100 protein, alpha chain.
GN S100A1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86248083; PubMed=3755105;
RA Kuwano R., Maeda T., Usui H., Araki K., Yamakuni T., Ohshima Y.,
RA Kurihara T., Takahashi Y.;
RT "Molecular cloning of cDNA of S100 alpha subunit mRNA.";
RL FEBS Lett. 202:97-101(1986).
RN [2]
RP SEQUENCE.
RX MEDLINE=81236562; PubMed=7250124;
RA Iisobe T., Okuyama T.;
RT "The amino-acid sequence of the alpha subunit in bovine brain S-100a
RT protein.";
RL Eur. J. Biochem. 116:79-86(1981).
RN [3]
RP METAL ION-BINDING PROPERTIES.
RX MEDLINE=84000339; PubMed=6615778;
RA Baudier J., Gerard D.;
RT "Ions binding to S100 proteins: structural changes induced by calcium
RT and zinc on S100a and S100b proteins.";
RL Biochemistry 22:3360-3369(1983).
CC -I- FUNCTION: WEAKLY BINDS CALCIUM BUT BINDS ZINC VERY TIGHTLY-
CC DISTINCT BINDING SITES WITH DIFFERENT AFFINITIES EXIST FOR BOTH

```

```

CC IONS ON EACH MONOMER. PHYSIOLOGICAL CONCENTRATIONS OF POTASSIUM
CC ION ANTAGONIZE THE BINDING OF BOTH DIVALENT CATIONS, ESPECIALLY
CC AFFECTING HIGH-AFFINITY CALCIUM-BINDING SITES.
CC -I- SUBUNIT: DIMER OF EITHER TWO ALPHA CHAINS, OR TWO BETA CHAINS, OR
CC ONE ALPHA AND ONE BETA CHAIN.
CC -I- TISSUE SPECIFICITY: ALTHOUGH PREDOMINANT AMONG THE WATER-SOLUBLE
CC BRAIN PROTEINS, S-100 IS ALSO FOUND IN A VARIETY OF OTHER TISSUES.
CC -I- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC -I- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
DR PIR; A24156; A24156.
DR HSSP; P04631; 1B4C.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001751; S100_CABP.
DR Pfam; PF00036; ehand; 1.
DR DR; PF01023; S_100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Calcium-binding; zinc; Metal-binding.
FT INIT_MET 0
FT MOD_RES 1 1
FT CA_BIND 19 32 EF_HAND 1 (LOW AFFINITY).
FT CA_BIND 62 73 EF_HAND 2 (HIGH AFFINITY).
FT CONFLICT 64 64 N -> D (IN REF. 2).
SQ SEQUENCE 93 AA; 10387 MW; 0457D0F44819B89B CRC64;

alignment_scores:
  Quality: 146.00      Length: 92
  Ratio: 2.246        Gaps: 2
  Percent Similarity: 70.652      Percent Identity: 38.043

alignment_block:
US-09-806-382A-1 x S10A_BOVIN ..

Align seg 1/1 to: S10A_BOVIN from: 1 to: 93

7 ACCGAGCTGGAGAAAGCCTTGAACCTATCATCGACGTCTACCAAGTA 56
  ::::::::::::::::::::|::::::::::::::::::|::
2 SerGluLeuGlnThrAlaMetGlnThrLeuIleAsnValPheHisAlaH1 18
  ::::::::::::::::::::|::::::::::::::::::|::
57 CTCCTCATTAAGGAGGAAATTTCCATGCCGCTACAGGATGACGTGAAGA 106
  ::::::::::::::::::::|::::::::::::::::::|::
18 sSerGlyLysGlnGlyAspLysTyrLysLeuSerLysLysGlnLeuLysG 35
  ::::::::::::::::::::|::::::::::::::::::|::
107 AATTGCTAGAGACCGAGTGTCTCTCAGTATATC...AGAAAAAGGTTGCA 153
  ::::::::::::::::::::|::::::::::::::::::|::
35 LuLeuLeuGlnThrGlnLeuSerGlyPheLeuAspAlaGlnLysAspAla 51
  ::::::::::::::::::::|::::::::::::::::::|::
154 GACGTC.....TGTTCAAGAGTTGATATCAACACTGATGTTGC 194
  ::::::::::::::::::::|::::::::::::::::::|::
52 AspAlaValAspLysValMetLysGlnLeuAspGlnAsnGlyAspGlyG1 68
  ::::::::::::::::::::|::::::::::::::::::|::
195 AGTTAACTTCCAGAGAGTTCTCTGATTAAGATGGCGGTGCAG 244
  ::::::::::::::::::::|::::::::::::::::::|::
68 uValAspPheGlnGlnPheValIleValIleValAlaIleuThrValAlaC 85
  ::::::::::::::::::::|::::::::::::::::::|::
245 CCCACAAAAAAGCCATGAAGAAAGC 270
  ::::::::::::::::::::|::::::::::::::::::|::
85 ysAsnAsnPhePheTrpGlnAsnSer 93
  ::::::::::::::::::::|::::::::::::::::::|::

seq_name: SwissProt_40:S10A_HUMAN
seq_documentation_block:
ID S10A_HUMAN STANDARD; PRT; 93 AA.
AC P23297;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-100 protein, alpha chain (S100 calcium-binding protein A1).
GN S100A1 OR S100A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

seq_name: swissprot_40:s10A_MOUSE

seq_documentation_block:
ID      S10A_MOUSE      STANDARD:      PRT:      93 AA.
AC      P56565; 088949;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      S-100 protein, alpha chain (S100 calcium-binding protein A1).
GN      S100A1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA      Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
RA      Gelsel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
RA      Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,
RA      Theising B., Wyllie T., Lennon G., Soares B., Wilson R., Waterston R.,
RL      Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99117144; PubMed=9920416;
RA      Rindner K., Ilg E.C., Niggli F.K., Heizmann C.W., Schaefer B.W.;
RL      "Clustered organization of S100 genes in human and mouse.";
RL      Biochim. Biophys. Acta 1448:254-263(1998).
CC      -1- FUNCTION: WEAKLY BINDS CALCIUM BUT BINDS ZINC VERY TIGHTLY.
CC      DISTINCT BINDING SITES WITH DIFFERENT AFFINITIES EXIST FOR BOTH
CC      IONS ON EACH MONOMER. PHYSIOLOGICAL CONCENTRATIONS OF POTASSIUM
CC      ION ANTAGONIZE THE BINDING OF BOTH DIVALENT CATIONS, ESPECIALLY
CC      AFFECTING HIGH-AFFINITY CALCIUM-BINDING SITES (BY SIMILARITY).
CC      -1- SUBUNIT: DIMER OF EITHER TWO ALPHA CHAINS, OR TWO BETA CHAINS, OR
CC      ONE ALPHA AND ONE BETA CHAIN (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC      -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; AA000715; -; NOT_ANNOTATED_CDS.
DR      EMBL; AA207749; -; NOT_ANNOTATED_CDS.
DR      EMBL; AA500563; -; NOT_ANNOTATED_CDS.
DR      EMBL; AAA32539; -; NOT_ANNOTATED_CDS.
DR      EMBL; AF087687; AAC64108.1; -.
DR      HSSP; P04631; 1B4C.
DR      MGD; MGI:1338917; S100a1.
DR      InterPro; IPR002048; EF-hand.
DR      InterPro; IPR001751; S100_CaBP.
DR      Pfam; PF00036; efhand; 1.
DR      Pfam; PF1023; S100; 1.
DR      PROSITE; PS00018; EF_HAND; 1.
DR      PROSITE; PS00303; S100_CaBP; 1.
DR      Calcium-binding; zinc; Metal-binding.
DR      INTT_MET 0
DR      CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).
DR      CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY) (BY

```





```

34  IuLeuIleAsnAsnGluLeuSerHisPheLeuGluGluIleLysGluGln 50
    ::|||:.....|||  :::::
145 AAGGTCGACAGCTCTGGTTCAAAGAGTTGATATCAACACTGATGTGC 194
    ::  ::|||  ::::: |||||  ::  |||||
51 GluValValAspLysValMetGluThrLeuAspSerAspGlyAspGlyG1 67
    ::|||:|||||:|||||:  ::|||
195 AGTTAACTTCACAGAGTTCTCATCTGTGATGAATAAGATGGCGTGCAG 244
    ::|||:|||||:|||||:  ::|||
67 uCysAspPheGlnGluPheMetAlaPheValAlaMetIleThrThrAlaC 84
    ::|||:|||||:|||||:  ::|||
245 CCCACAAA 252
84 ySHISGlu 86
```

OM of: US-09-806-382a-1 to: SPREMBL\_19:\* out\_format : pfs

Date: Sep 9, 2002 3:24 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

#### Command line parameters:

-MODEL=frame\_n2p.model -DEV=xlp  
-O=/cg2\_1/USPRO.spool/US09806382/runatc\_09092002\_143849\_15535/app-query.fasta\_1.752  
-DB=SPREMBL\_19 -OFMT=fasta -SUFFIX=rsp -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -RGAPEXT=7.000 -YGAPO=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsumb2  
-TRANS=human40.csl -LIST=45 -DGCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09806382.@CGN1.1.279 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPPY -WAIT -THREADS=1

#### Search information block:

Query: US-09-806-382a-1  
Query length: 282  
Database: SPREMBL\_19:\*  
Database sequences: 56222  
Database length: 172994929  
Search time (sec): 104.890000

#### Score list:

Sequence	Strd	Orig	ZScore	Escore	Len	Documentation
SP_mammal:Q28714	+	269.50	584.94	4.3e-24	73	Q28714 oryctolagus cuniculus (rat)
SP_vertibrate:093395	+	146.00	312.10	2.6e-08	101	Q93395 salvelinus fontinalis (fish)
SP_rident:09D3M4	+	140.00	299.63	2.6e-08	94	Q9D3M4 mus musculus (mouse). s10
SP_rident:091V77	+	140.00	299.63	2.6e-08	94	Q91V77 m. 11 days embryo cdna, rat
SP_human:09H4U0	+	139.00	308.64	2.8e-08	27	Q9H4U0 homo sapiens (human). cal
SP_rident:0925J3	+	139.00	297.63	3.4e-08	92	Q925J3 cricetus griseus (chinese)
SP_rident:09JL08	+	135.00	290.26	1.0e-07	79	Q9JL08 mus musculus (mouse). s10
SP_mammal:09TV56	+	127.00	270.57	1.0e-06	101	Q9TV56 canis familiaris (dog). s10
SP_human:09H8G3	+	126.00	254.12	1.5e-06	495	Q9H8G3 homo sapiens (human). tu
SP_human:09H8Y0	+	125.50	266.20	1.5e-06	114	Q9H8Y0 homo sapiens (human). ml
SP_human:09H4U0	+	125.00	278.73	1.4e-06	25	Q9H4U0 homo sapiens (human). cal
SP_human:09H4U0	+	122.00	252.94	4.5e-06	213	Q9H4U0 homo sapiens (human). cal
SP_human:09H8G3	+	118.50	232.26	1.1e-05	98	Q9H8G3 homo sapiens (human). s10
SP_human:001720	+	114.50	227.39	4.3e-05	591	Q01720 homo sapiens (human). fi
SP_human:09H4U0	+	114.50	226.04	4.4e-05	687	Q9H4U0 homo sapiens (human). d
SP_human:005331	+	114.00	219.81	5.6e-05	1218	Q05331 homo sapiens (human). d
SP_vertibrate:09P5F6	+	112.00	238.62	6.6e-05	92	Q9P5F6 ictalurus punctatus (chinese)
SP_rident:09R2B7	+	109.00	232.36	0.0002	89	Q9R2B7 rattus norvegicus (rat). s10
SP_mammal:09R1R6	+	107.00	230.15	0.0003	70	Q9R1R6 bos taurus (bovine). com
SP_human:09H8G3	+	105.50	223.40	0.0004	103	Q9H8G3 homo sapiens (human). s10
SP_rident:091X55	+	105.00	224.58	0.0005	80	Q91X55 mus musculus (mouse). s10
SP_human:09H8G3	+	100.00	226.09	0.0015	20	Q9H8G3 homo sapiens (human). cal
SP_rident:09D708	+	100.00	209.72	0.0020	124	Q9D708 mus musculus (mouse). s10
SP_vertibrate:09YH57	+	98.50	201.41	0.0033	217	Q9YH57 rana catesbeiana (bullfrog)
SP_human:09H8G3	+	97.50	207.23	0.0038	89	Q9H8G3 homo sapiens (human). s10
SP_rident:09H8G3	+	96.00	213.71	0.0049	30	Q9H8G3 homo sapiens (human). s10
SP_human:09H8G3	+	91.00	206.42	0.0188	20	Q9H8G3 homo sapiens (human). cal
SP_rident:09D3P1	+	88.00	168.78	0.0736	638	Q9D3P1 mus musculus (mouse). s10
SP_plant:09S531	+	84.50	171.15	0.1659	209	Q9S531 arabidopsis thaliana (maize)
SP_invertebrate:077372	+	84.50	142.69	0.2680	4981	Q77372 plasmodium falciparum (malaria)
SP_plant:09IKN1	+	83.50	160.26	0.2543	551	Q9IKN1 gossypium hirsutum (upland cotton)
SP_rident:09H8G3	+	81.50	112.86	0.3346	83	Q9H8G3 rattus norvegicus (rat). s10
SP_plant:080700	+	81.00	154.77	0.5128	553	Q80700 arabidopsis thaliana (maize)
SP_plant:09S5W7	+	79.50	152.25	0.7708	508	Q9S5W7 gossypium arboreum (cotton)
SP_bacteriap:09H8R19	+	74.50	154.35	2.51	119	Q9H8R19 mycoplasma pulmonis (hygiene)
SP_invertebrate:09VBL2	+	74.50	149.42	2.73	206	Q9VBL2 drosophila melanogaster
SP_invertebrate:09NAY9	+	74.00	149.06	3.10	190	Q9NAY9 naegleria fowleri (amoeba)
SP_invertebrate:0188R7	+	74.00	144.38	3.36	320	Q188R7 caenorhabditis elegans
SP_plant:09SDM4	+	74.00	138.81	3.69	595	Q9SDM4 dunaliella tertiolecta
SP_plant:09STR5	+	74.00	135.10	3.93	900	Q9STR5 arabidopsis thaliana (maize)

SP\_human:09H8G3 + 73.50 143.46 3.85 314 | Q9H8G3 homo sapiens (human).  
SP\_invertebrate:09VZ14 + 73.50 140.76 4.03 424 | Q9VZ14 drosophila melanogaster  
SP\_invertebrate:09YXU1 + 73.00 129.65 5.49 1294 | Q9YXU1 drosophila melanogaster  
SP\_invertebrate:0960D0 + 73.00 128.96 5.56 1398 | Q960D0 drosophila melanogaster  
SP\_plant:09S168 + 72.50 125.63 6.64 1794 | Q9S168 arabidopsis thaliana

seq\_name: SP\_mammal:Q28714

#### seq\_documentation\_block:

ID	Q28714	PRELIMINARY:	PRT:	73 AA.
AC	Q28714:			
DT	01-NOV-1996 (TEMBLrel. 01, Created)			
DT	01-NOV-1996 (TEMBLrel. 01, Last sequence update)			
DE	01-DEC-2001 (TEMBLrel. 19, Last annotation update)			
DE	MACROPHAGE MIGRATION INHIBITORY FACTOR-RELATED PROTEIN-8 (FRAGMENT).			
GN	MRP-8.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN=NEW ZEALAND WHITE.			
RX	MEDLINE=94198229; PubMed=8148323;			
RA	Mori S., Goto K., Goto F., Murakami K., Ohkawara S., Yoshinaga M.;			
RT	"Dynamic changes in mRNA expression of neutrophils during the course			
RT	of acute inflammation in rabbits."			
RL	Int. Immunol. 6:149-156(1994).			
CC	-1- SIMILARITY: BELONGS TO THE S-100 FAMILY.			
EMBL	D17405; BAA04228.1; -.			
DR	HSSP; P02632; ICB1.			
DR	InterPro; IPR002048; EF-hand.			
DR	InterPro; IPR001751; S100_CaBP.			
DR	Pfam; PF01023; S_100; 1.			
FT	NON_TER			
SQ	SEQUENCE 73 AA; 8381 MW; F69089226601A163 CRC64;			

#### alignment\_scores:

Quality: 269.50 Length: 74  
Ratio: 4.278 Gaps: 1  
Percent Similarity: 85.135 Percent Identity: 74.324

#### alignment\_block:

US-09-806-382a-1 x Q28714

Align seg 1/1 to: Q28714 from: 1 to: 73

```
1 ATGTTGACGACGTCGAGAAAGCTTAACCTATCATGACGCTACCA 50
|||||
1 MetProthraspleuGlusSerleuAsnSerlellepheValTyrH1 17
51 CAATACCTCCCTGATTAAGGGAATTCATGCCGCTACAGGAGGACC 100
|||||
17 SlvGcySerleuGlulysGlnTyrGlnAlaLeuTyr...GlyAsp 33
101 TGAAGAAATGCTAGACGAGTGTCTCAGTATATCAGAAAGGCT 150
|||||
33 eulysleuLeuAlaThrGlnCysProGlnTyrSerleYslvslvAsp 49
151 GCAGACGTCGCTCAAGAGCTTGATATACACAGCTGAGCTCACTTAA 200
|||||
50 AlaspserrTprheYslvslvAspIleasSerAspIlylaseras 66
201 CTTTCAGAGTCTCCTCATCTCG 222
|||||
66 nphcGlnGluPheleuIleu 73
seq_name: SP_vertibrate:093395
seq_documentation_block:
ID Q93395 PRELIMINARY: PRT: 101 AA.
```

AC 093395;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE S100-LIKE CALCIUM BINDING PROTEIN.  
 GN S100.  
 OS Salvelinus fontinalis (Brook trout) (Brook char).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.  
 OC NCBI\_TaxID=8038;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=OVULATORY OVARY;  
 RX MEDLINE=20534789; PubMed=11080585;  
 RA Bohe J., Goetz F.W.;  
 RT "A S100 homologue mRNA isolated by differential display PCR is down-  
 regulated in the brook trout (Salvelinus fontinalis) post-ovulatory  
 RT ovary";  
 RL Gene 257,187-194(2000).  
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
 DR EMBL; AF077613; AAC28367.1; -  
 DR HSSP; P02633; 4ICB.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR001751; S100\_CABP.  
 DR Pfam; PF00036; ehand.1.  
 DR Pfam; PF01023; S\_100; 1.  
 DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_1.  
 DR PROSITE; PS00303; S100\_CABP; 1.  
 KW Calcium-binding.  
 SQ SEQUENCE 101 AA; 11285 MW; BF45582FF9279DOA CRC64;

alignment\_scores:  
 Quality: 146.00 Length: 88  
 Ratio: 2.179 Gaps: 1  
 Percent Similarity: 76.136 Percent Identity: 31.818

alignment\_block:

US-09-806-382a-1 x 093395 ..

Align seg 1/1 to: 093395 from: 1 to: 101

```

1 ATGTGACCGAGCTGGAGAAAGCCTGAACCTATCATGACGCTACCA 50
|||||.....|.....|.....|.....|.....|.....|.....|.....|
1 MetGlySerGlnLeuGlnSerMetGlnSerLeuIleThrValPheH 17
51 CAAGTACTCCCTGATTAAGGGAATTTCCATGCCGTCTACAGGAGTAC 100
|.....|.....|.....|.....|.....|.....|.....|.....|
17 sarGlyrAlaAspLysAspLysAspLysAsnThrLeuSerLysLysL 34
101 TGAAGAAATTTGCTAGACAGCCGAGTGTCTCAGTATATCAGAAAA 147
|||||.....|.....|.....|.....|.....|.....|.....|.....|
34 eulysGlnLeuMetGlnThrGlnLeuAlaSerPheLeuLysSerGln 50
148 .....GGTGACAGCTGCTGATCAAGAGTTGATATCAACACTGA 188
|||||.....|.....|.....|.....|.....|.....|.....|.....|
51 AspProAlaIleAlaIleAspThrIleMetLysAspLeuAspGlnsn 67
189 TGTGACAGTTAACTTCAGAGAGTTCCTCATTCGTGATTAAGATGAG 238
|||||.....|.....|.....|.....|.....|.....|.....|.....|
67 pGlyLysValSerPheGlnGlnPheValSerLeuValValGlyLeuS 84
239 TGGCAGCCCAAAA 252
:::|.....|.....|
84 leAlaCysGln 88

```

seq\_name: sp\_rodent:Q9D3M4

seq\_documentation\_block:  
 ID Q9D3M4  
 AC Q9D3M4; BRELIMINARY; PRT; 94 AA.

DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE S100 CALCIUM BINDING PROTEIN A1.  
 GN S100A1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=HEAD;  
 RX MEDLINE=21085660; PubMed=1121851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G.,  
 RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Winkler L.,  
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawai J., Koltzski S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
 DR EMBL; AK017279; BAB30670.1; -  
 DR HSSP; P04631; 1B4C.  
 DR MGD; MGI:1336917; S100a1.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR001751; S100\_Cabp.  
 DR Pfam; PF01023; S\_100; 1.  
 DR SMART; SM00054; EFh; 1.  
 DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_1.  
 SQ SEQUENCE 94 AA; 10487 MW; 7C8EB817D135C2D15 CRC64;

alignment\_scores:  
 Quality: 140.00 Length: 94  
 Ratio: 2.154 Gaps: 2  
 Percent Similarity: 69.149 Percent Identity: 37.234

alignment\_block:

US-09-806-382a-1 x Q9D3M4 ..

Align seg 1/1 to: Q9D3M4 from: 1 to: 94

```

1 ATGTGACCGAGCTGGAGAAAGCCTGAACCTATCATGACGCTACCA 50
|||||.....|.....|.....|.....|.....|.....|.....|.....|
1 MetGlySerGlnLeuGlnSerMetGlnSerLeuIleThrValPheH 17
51 CAAGTACTCCCTGATTAAGGGAATTTCCATGCCGTCTACAGGAGTAC 100
|.....|.....|.....|.....|.....|.....|.....|.....|
17 salanHsSerGlyLysGlnGlnLysAspLysLysSerLysLysL 34
101 TGAAGAAATTTGCTAGACAGCCGAGTGTCTCAGTATATCAG...AAAAG 147
|||||.....|.....|.....|.....|.....|.....|.....|.....|
34 eulysAspLeuLeuGlnThrGlnLeuSerGlyPheLeuAspValGln 50
148 GGTGACAGCTC.....TGGTTCAAGAGCTGATATCAACACTGA 188
|||||.....|.....|.....|.....|.....|.....|.....|.....|
51 AspAlaAspAlaValAspLysValIleLysGlnLeuAspLysn 67
189 TGTGACAGTTAACTTCAGAGAGTTCCTCATTCGTGATTAAGATGAG 238
|||||.....|.....|.....|.....|.....|.....|.....|.....|

```

```

67 pglgluValAspPheLysGluTyrValValLeuValAlaAlaLeuThrV 84
239 TGGCAGCCACACAAAAAGCCATGAAGAAGC 270
|||||:|||||
84 aAlaLacYsAsnAspPheTrpGluThrSer 94

seq_name: sp_rodent:Q91V77

seq_documentation_block:
ID Q91V77 PRELIMINARY; PRT; 94 AA.
AC Q91V77:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 11 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2700088D09, FULL INSERT SEQUENCE (S100 CALCIUM BINDING PROTEIN
DE A1) (ADULT MALE KIDNEY CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:0610031F03, FULL INSERT SEQUENCE).
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA Arai K., Aizawa T., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara K., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imetani K., Ishii Y., Itoh M., Izawa M., Kato H., Kawai J., Kojima Y.,
RA Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T.,
RA Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D.,
RA Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y.,
RA Suzuki H., Tagami M., Tanigawa A., Takahashi F., Tanaka T., Tejima Y.,
RA Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=9279253; PubMed=10349636;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN 15
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsuno H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format

```

```

RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN 16
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN 17
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Du X.-J., Cole T.J., Tennis N., Gao X.-M., Kontgen F., Kemp B.E.,
RA Heierhorst J.;
RT "Impaired cardiac contractility without cardiomyopathy in S100A1-
RT deficient mice.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK012578; BAB28330.1; -
DR EMBL; BC005590; AAH05590.1; -
DR EMBL; AF368423; AAL14436.1; -
DR EMBL; AK002721; BAB23308.1; -
SQ SEQUENCE 94 AA; 10505 MW; 7C8E817BBF36ED15 CRC64;

alignment_scores:
Quality: 140.00 Length: 94
Ratio: 2.154 Gaps: 2
Percent Similarity: 69.149 Percent Identity: 37.234

alignment_block:
US-09-806-382A-1 x Q91V77 ..
Align seg 1/1 to: Q91V77 from: 1 to: 94

1 ATGTGACCGAGCTCGAGAAAGCCTTGACTCATCATGACGCTTACCA 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MetGluSerGluLeuGluSerAlaMetGluThrLeuLeuAlaPheH 17
51 CAGTACTCCCTGATAAGGGAATTCCATGCCGCTACAGAGATGACC 100
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17 SAlaHISerG|LylYsGluGlnLysPlYsTrYlYsLeuSerLylYsGlu 34
101 TGAAGAAATGCTAGACGACGAGTCTCCTCAGTATATACAG...AAAAG 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 euLysAspLeuLeuGlnThrGluLeuSerGlyPheLeuAspValGlnLys 50
148 GGTGCACAGCTC.....TGGTCAAGAAGTTGGATATCAACACTGA 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||
51 AspAlaAspAlaValAspLysValMetLysGluLeuAspLysnLys 67
189 TGGTGCAGTTAAGTCCAGAGAGTTCCTCATTCGTGATTAAGATGGCG 238
|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 pglgluValAspPheLysGluTyrValValLeuValAlaAlaLeuThrV 84
239 TGGCAGCCACACAAAAAGCCATGAAGAAGC 270
|||||:|||||
84 aAlaLacYsAsnAspPheTrpGluThrSer 94

seq_name: sp_human:Q9UCJ0

seq_documentation_block:
ID Q9UCJ0 PRELIMINARY; PRT; 27 AA.
AC Q9UCJ0:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CALPROTECTIN LIL SUBUNIT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE.
RX MEDLINE=93139333; PubMed=8423249;
RA Miyasaki K.T., Bodeau A.L., Murthy A.R., Lehrer R.I.;
RT "In vitro antimicrobial activity of the human neutrophil cytosolic S-

```

RT 100 protein complex, calprotectin, against Capnocytophaga sputigena."  
 RL J. Dent. Res. 72:517-523(1993).  
 SO SEQUENCE 27 AA; 3178 MW; BC5E928BE805CE8 CRC64;

## alignment\_scores:

Quality: 139.00 Length: 27  
 Ratio: 5.148 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-806-382A-1 x Q9UCJ0 ..

Align seg 1/1 to: Q9UCJ0 from: 1 to: 27

```

1  ANGTGACCGAGCTGGAGAAACCTTGAACTATCATCGACGTACCA 50
  |||||||
1  MetLeuThrgluLeuGluLysAlaLeuAsnSerIleIleAspValThr 17
51 CAAGTACTCCCTGATTAAGGGAATTTCCAT 81
  |||||||
17 slystySerIleuIleLysGlyAsnPhen 27

```

seq\_name: sp\_rhodent:Q925T3

## seq\_documentation\_block:

ID Q925T3 PRELIMINARY; PRT; 92 AA.

AC Q925T3;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE S100B.  
 OS Cricetus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kuge O., Yamakawa Y., Nishijima M.;  
 RT "Enhancement of transport-dependent decarboxylation of  
 RT phosphatidylserine by S100B protein in permeabilized Chinese hamster  
 RT ovary cells.";  
 RL J. Biol. Chem. 0:0-0(2001).  
 DR EMBL: AB056121; BAB43945.1;  
 SO SEQUENCE 92 AA; 10749 MW; AFS0107EC2BEDF6B CRC64;

## alignment\_scores:

Quality: 139.00 Length: 92  
 Ratio: 2.106 Gaps: 1  
 Percent Similarity: 71.739 Percent Identity: 31.522

## alignment\_block:

US-09-806-382A-1 x Q925T3 ..

Align seg 1/1 to: Q925T3 from: 1 to: 92

```

4  TTGACCGAGCTGGAGAAACCTTGAACTATCATCGACGTACCA 53
  |||||||
1  MetSerGluLeuGluLysAlaMetValAlaLeuIleAspIlePheHis 17
54 GTATCCCTGATTAAGGGAATTTCCATCCCTACAGGGATGACCTGA 103
  |||||||
17 nTysSerIlyArgGluGlyAspLysHisLysLeuLysSerGluLeu 34
104 AGAATTCTAGAGACCGAGCTGCTCCTCAGTATATC.....AGG 141
  |||||||
34 ysgLleuIleAsnAsnGluLeuSerHisPheLeuGluGluLysGlu 50
142 AAAAAGGTGCAGACGTCTGGTTCAAGAAGTTGGATATCAACACTG 191
  |||||||
51 GIngluValIleAspLysValMetGluThrLeuAspGluAspGlyAsp 67

```

```

192 TGCAGTTACTCCAGAGTTCCTCATTCGGTATGAAGATGGCGTGG 241
  |||||||
67 ygluLysAspPheGluGluPheMetAlaPheValSerMetValThr 84

```

242 CAGCCCAACAAAAAGCCATGAAGAA 267

```

  |||||||
84 lacysHisGluPhePheGluGluGlu 92

```

seq\_name: sp\_rhodent:Q9JL08

## seq\_documentation\_block:

ID Q9JL08 PRELIMINARY; PRT; 79 AA.

AC Q9JL08;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE S100 CALCIUM BINDING PROTEIN A1 (FRAGMENT).  
 GN S100A1 OR S100A1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kiewitz R., Lyons G.E., Schafer B.W., Heizmann C.W.;  
 RT "Transcriptional regulation of S100A1 and expression during mouse  
 RT heart development.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
 DR EMBL: AF218353; AAF32320.1; -.  
 DR HSSP: P04631; 1B4C.  
 DR MGD: MGI:1338917; S100a1.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR001751; S100\_Cabp.  
 DR Pfam: PF01023; S\_100; 1.  
 DR PROSITE: PS00018; EF\_HAND; UNKNOWN\_1.  
 DR PROSITE: PS00303; S100\_CABP; 1.  
 FT NON-TER 79  
 SO SEQUENCE 79 AA; 8863 MW; F94EDA3A798815D5 CRC64;

## alignment\_scores:

Quality: 135.00 Length: 79  
 Ratio: 2.368 Gaps: 2  
 Percent Similarity: 72.152 Percent Identity: 39.241

## alignment\_block:

US-09-806-382A-1 x Q9JL08 ..

Align seg 1/1 to: Q9JL08 from: 1 to: 79

```

1  ANGTGACCGAGCTGGAGAAACCTTGAACTATCATCGACGTACCA 50
  |||
1  MetGlySerGluLeuGluSerAlaMetGluThrLeuIleAsnValPhe 17
51 CAAGTACTCCCTGATTAAGGGAATTTCCATCCCTACAGGGATGAC 100
  |||||||
17 salahisSerAlaGluGluGlyAspLysThrLysLeuSerLysGlu 34
101 TGAAGAAATTCCTAGAGACCGAGCTGCTCAGTATATCAGG...AAAAG 147
  |||||||
34 eulLysAspLeuLeuGlnThrGluLeuSerGlyPheLeuAspValGln 50
148 GGTGCAGAGTCT.....TGTTCAAGAAGTTGGATATCAACACTGA 188
  |||||||
51 AspAlaAspAlaValAspLysValMetLysGluLeuAspGluAsnGly 67
189 TGGTGCAGTTACTTCAGAGACTTCTCATTTCTGTG 225
  |||
67 pGlyGluValAspPheLysGluThrValValLeuVal 79

```

seq\_name: sp\_mammal:Q9TV56

```

seq_documentation_block:
ID Q9TV56 PRELIMINARY; PRT; 101 AA.
AC Q9TV56;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE METASTASIN.
GN MTS1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MADIN-DURBY; TISSUE=MADIN DARRY CANINE KIDNEY (MDCK);
RA Miyamoto H., Hasegawa K., Kim K., Sato H.;
RT "Expression of metastasis associated mts1 gene is co-induced with
RT membrane type-1 matrix metalloproteinase (MT1-MMP) during oncogenic
RT transformation and tubular formation of madin darry canine kidney
RT (MDCK) epithelial cells."
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL: AB031064; BAA83419.1; -.
DR HSSP: P30801; 1A03.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_CaBP.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF01023; S_100; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE: PS00303; S100_CaBP; 1.
KW Calcium-binding.
SQ SEQUENCE 101 AA; 11847 MW; 5535387DB7577DF0 CRC64;

alignment_scores:
Quality: 127.00 Length: 79
Ratio: 2.352 Gaps: 1
Percent Similarity: 68.354 Percent Identity: 35.443

alignment_block:
US-09-806-382a-1 x Q9TV56 ..
Align seg 1/1 to: Q9TV56 from: 1 to: 101

1 ATGTTGACGAGCTGGAGAAAGCCTGAACCTATCATCATGACGCTACCA 50
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MetThrPheProLeuGluLysAlaLeuAspValMetValSerThrPheH 17
51 CAGGTACTCCCTGATTAAGGGGAAATTCATGCCGCTACAGGAGATGACC 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 slvstYrSerGlyLysGluGlyAspLysPheLysLeuAsnArgSerGluL 34
101 TGAAGAAATTTGTAAGACGAGCTGCTCAGATATATCAGAAAGAGGT 150
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 euLysLeuLeuLeuMetArgGluLeuProSerPheLeuGlyLysArgThr 50
151 GCAGAGCTGTGTTCAAGAG.....TTGATATCAACACTGA 188
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 AspGluAlaAlaPheGlnLysLeuMetSerAsnLeuAspSerAsnArgAs 67
189 TGGTGCAGTTAACTTCAGAGAGTTCCTCATTCGTGTG 225
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 pasngluValAspPheGlnGluTyrCysValPheLeu 79

seq_name: sp_human:Q9UBG3

seq_documentation_block:
ID Q9UBG3 PRELIMINARY; PRT; 495 AA.
AC Q9UBG3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

```

```

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TUMOR RELATED PROTEIN.
GN DRCL OR PDRL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Xu Z.X., Wang M.R., Cai Y., Xu X., Han Y.L., Wang X.Q., Wu M.;
RT "Cloning and characterization of a novel gene associated with human
RT cancer."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Xu Z.X., Wang M.R., Cai Y., Xu X., Han Y.L., Wang X.Q., Wu M.;
RT "Cloning and characterization of a novel CDNA associated with human
RT cancer."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL: AF07831; AAD55747.1; -.
DR HSSP: P02633; 3ICB.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_CaBP.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF01023; S_100; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
KW Calcium-binding.
SQ SEQUENCE 495 AA; 53533 MW; C4882A11B4E64DC3 CRC64;

alignment_scores:
Quality: 126.00 Length: 93
Ratio: 2.032 Gaps: 1
Percent Similarity: 66.667 Percent Identity: 34.409

alignment_block:
US-09-806-382a-1 x Q9UBG3 ..
Align seg 1/1 to: Q9UBG3 from: 1 to: 495

4 TTGACGAGCTGGAGAAAGCCTGAACCTATCATCATGACGCTACCAAA 53
::: ||||| ::: ||||| ||||| ||||| ||||| ||||| |||||
1 MetProGlnLeuGlnAsnIleAsnGlyIleIleGluAlaPheArgAr 17
54 GTACGCCCTGATTAAGGGGAAATTCATGCCGCTACAGGAGATGACCTGA 103
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 gtyrAlaArgThrGluGlyAsnCysThrAlaLeuThrArgGlyLeuLeuL 34
104 AGAATTTGTAAGACGAGCTGCTCAGATATATCAGAAAGAGGT..... 144
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 ysArgLeuLeuGlnGlnGluPheAlaAspValIleValLysProHisAsp 50
145 ..AAGGTCGACAGCTGTGTTCAAGAGCTTGATATCAACACTGAG 191
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 ProAlaThrValAspGluValLeuArgLeuLeuAspGluAspHisThrG 67
192 TGCAGTTAACTTCAGAGAGTTCCTCATTCGTGATTAAGATGGCGCTGG 241
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 yThrValGluPheLysGluPheLeuValLeuValPheLysValAlaGlnA 84
242 CAGCCACAAAAAGCCATGAGAAAGC 270
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 lacysPheLysThrLeuSerGluSerAla 93

seq_name: sp_human:Q9NTW0

seq_documentation_block:
ID Q9NTW0 PRELIMINARY; PRT; 114 AA.
AC Q9NTW0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)

```

DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
DE MIGRATION INHIBITOR FACTOR-RELATED PROTEIN 14 VARIANT E.  
GN S100A9.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang M., Xu X., Cai Y., Xu H., Han Y., Xu Z., Wu M.,  
RT "Human gene for migration inhibitory factor-related protein 14  
RT (MRP14), variant allele."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
DR EMBL; AF237582; AAF62537.1; -.  
DR HSSP; P02638; ICFP.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR001751; S100\_Cabp.  
DR Pfam; PF01023; S\_100; 1.  
DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_1.  
DR PROSITE; PS00303; S100\_CABP; 1.  
SQ SEQUENCE 114 AA; 13261 MW; D701528635F4DEC2 CRC64;

alignment\_scores:  
Quality: 125.50 Length: 92  
Ratio: 1.902 Gaps: 1  
Percent Similarity: 71.739 Percent Identity: 26.087

alignment\_block:  
US-09-806-382a-1 x Q9NYW0 ..

Align seg 1/1 to: Q9NYW0 from: 1 to: 114

```
4 TTGACCGAGCTGGAGAAAGCCTTGAACCTATCATCGAGCTACACCAA 53
   :::::::::::::::::::: ::::::::::::::::::::
5 MetSerGlnLeuGlnArgAsnIleGluThrIleAsnThrPheArg1 21
54 GATCTCCCTGATTAAGGGGAATTCATGCCGCTACAGGAGGATGACTGA 103
   :::::::::::::::::::: ::::::::::::::::::::
21 nYrSerValIlysLeuGlyHisProAspThrLeuAsnGlnGlyLuphel 38
104 AGAAATTGCTAGAGACCGAGTGTCTCACTATATCAGAAA..... 144
   :::::::::::::::::::: ::::::::::::::::::::
38 ygluLeuValArgLysAspLeuGlnAsnPhenLeuLysGluAsnLys 54
145 .....AAGGTCAGAGCTGTGGTTCAAAGAGTTGGATATCAACACTGA 188
   :::::::::::::::::::: ::::::::::::::::::::
55 AsnGluLysValIleGluHisIleMetGluAspLeuAspThrAsnAlaAs 71
189 TGGTGCAGTACTTCAGAGAGTTCCTCATTCCTGGTGAATAAAGATGGCG 238
   :::::::::::::::::::: ::::::::::::::::::::
71 pLysGlnLeuSerPheGlnGluPheIleMetLeuMetAlaArgLeuThr 88
239 TGGCAGCCCAACAAAAAGCCATGAA 264
   :::::::::::::::::::: ::::::::::::::::::::
88 rPalaserHisGlnLysMetHisGlu 96
```

seq\_name: sp\_human:Q9NYW0

seq\_documentation\_block:  
ID Q9NYW0 PRELIMINARY; PRT; 25 AA.

AC Q9NYW0:  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)  
DE CALCIUM-BINDING MYELOID RELATED PROTEIN MRP-8 HOMOLOG (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE.  
RX MEDLINE=92406885; PubMed=1326551;  
RA Lemarchand P., Vaglio M., Manuel J., Markert M.;  
RT "Translocation of a small cytosolic calcium-binding protein (MRP-8) to  
RT plasma membrane correlates with human neutrophil activation."  
RL J. Biol. Chem. 267:19379-19382(1992).  
SQ SEQUENCE 25 AA; 2893 MW; 028BEC805CE88619 CRC64;

alignment\_scores:  
Quality: 125.00 Length: 25  
Ratio: 5.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-806-382a-1 x Q9NYW0 ..

Align seg 1/1 to: Q9NYW0 from: 1 to: 25

```
1 ATGTTGACCGAGCTGGAGAAAGCCTTGAACCTATCATCGAGCTACCA 50
   :::::::::::::::::::: ::::::::::::::::::::
1 MetLeuThrGlnLeuGlnLysAlaLeuAsnSerIleIleAspValTyrH1 17
51 CAAGTACTCCCTGATTAAGGGGAAT 75
   :::::::::::::::::::: ::::::::::::::::::::
17 sLysTYrSerLeuIleLysGlyAsn 25
```

seq\_name: sp\_human:Q9H4U1

seq\_documentation\_block:  
ID Q9H4U1 PRELIMINARY; PRT; 213 AA.

AC Q9H4U1:  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
DE DJ14N1.2 (NOVEL S-100/ICABP TYPE CALCIUM BINDING DOMAIN PROTEIN,  
DE SIMILAR TO TRICHOHYALIN) (FRAGMENT).  
GN DJ14N1.2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Laird G.;  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.  
DR EMBL; AL356504; CAC13173.1; -.  
DR HSSP; P02638; ICFP.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR001751; S100\_Cabp.  
DR Pfam; PF01023; S\_100; 1.  
DR SMART; SM00054; EFh; 1.  
DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_1.  
DR NON\_TER 213  
SQ SEQUENCE 213 AA; 24340 MW; B8C6E0810098E7D2 CRC64;

alignment\_scores:  
Quality: 122.00 Length: 107  
Ratio: 1.848 Gaps: 2  
Percent Similarity: 61.682 Percent Identity: 27.103

alignment\_block:  
US-09-806-382a-1 x Q9H4U1 ..

Align seg 1/1 to: Q9H4U1 from: 1 to: 213

```
4 TTGACCGAGCTGGAGAAAGCCTTGAACCTATCATCGAGCTACACCAA 53
   :::::::::::::::::::: ::::::::::::::::::::
1 MetThrAspLeuArgSerValIleThrValIleAspValPheTyrLys 17
54 GATCTCCCTGATTAAGGGGAATTCATGCCGCTACAGGAGGATGACTGA 103
```



```

|||||:|||||:|||||:|||||:|||||:
17 STYTHrlyslnaspglygIucYsglyTHrLeuSerlysglyGluLeuL 34
104 AGAAATTGCTAGAGACCGAGTGTCTCATGATATACAGAAAAAGGT... 150
34 YsgluLeuLeuGluLysGluLeuHisProValLeuLysAsnProAsp 50
151 .....GCAGACGTCTGGTTCAGAAAGTGGATATCAACACTGATGG 191
51 ProAspTHrValAspValIleMetHisMetLeuAspArgAspHisAsp 67
192 TGCAGTTAACTCCAGAGTCCATCTTCGATTCGGATAAACATGGCGTGG 241
67 gArgLeuAspPheThrGluPheLeuLeuMetIlePheLysLeuThrMet 84
242 CAGCCACACAAA.....AAA 255
84 lacysAsnLysValLeuSerLysGluTrcYrsLysAlaSerGlySerLys 100
256 AGCATTAGAAAGCCACAAA 276
101 LysHisArgArgGlyHisArg 107

```

seq\_name: sp\_human:Q9BU83

```

seq_documentation_block:
ID Q9BU83 PRELIMINARY; PRT; 98 AA.
AC Q9BU83;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE S100 CALCIUM-BINDING PROTEIN A2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-OVARY, ADENOCARCINOMA;
RA Strussberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC EMBL; BC002829; AA02829.1; -.
DR HSSP; P30801; IAO3.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001751; S100_CABP.
DR Pfam; PF01023; S_100; 1.
DR SMART; SM00054; EFh; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00303; S100_CABP; 1.
SQ SEQUENCE 98 AA; 1117 MW; 56D09548450142A9 CRC64;

```

alignment\_scores:

Quality:	118.50	Length:	85
Ratio:	2.194	Gaps:	2
Percent Similarity:	63.529	Percent Identity:	34.118

alignment\_block:

US-09-806-382a-1 x Q9BU83 ..

Align seg 1/1 to: Q9BU83 from: 1 to: 98

```

1  AAGTTGACGAGCTGGAGAAAGCTTGAACCTATCATCATGAGCTACCA 50
||| :|||:|||||:|||||:|||||:
2  MetCysSerSerLeuGluLysAlaLeuAlaValLeuValThrPheH 18
51  CAAGTACTCCCTGATTAAGGGAATTCATCCGCTACAGGAGATGACC 100
|||:|||||:|||||:|||||:|||||:
18  stYsYrSerCysGlnGluGlyAspLysPheLysLeuSerLysGlyL 35
101 TGAGAGAAATGCTAGAGACCGAGTGTCTCATATATACGAAAAAGGT 150
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

35  eLysGluLeuLeuHisLysGluLeuProSerPheValGlyLysVal 51
151 GCAGACGTCTGTTCAAGAG.....TTGCATATCAACACTGA 188
::: :|||:|||||:|||||:|||||:
52  AspGluGluGlyLeuLysLysLeuMetGlySerLeuAspGluAsnSer 68
189 TGGTGCAGTTAACTCCAGAG.....TTCCCATCTCGGAGATTA 229
| :|||:|||||:|||||:|||||:
68  pGlnGlnValAspPheGlnGluTrAlaValPheLeuAlaLeuIleThr 85
230 AGATG 234
|||
85  alMet 86

```

seq\_name: sp\_human:Q01720

```

seq_documentation_block:
ID Q01720 PRELIMINARY; PRT; 591 AA.
AC Q01720;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE FILAGGRIN PRECURSOR (PROFILAGGRIN) (FRAGMENT).
GN Fig.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RX MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Mirnusskiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus.";
RL J. Biol. Chem. 267:23772-23781(1992).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; L01089; AAA60177.1; -.
DR EMBL; L01090; AAA60176.1; -.
DR HSSP; P02593; ICDM.
DR MIM; 135940; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR003303; Filaggrin.
DR InterPro: IPR001751; S100_CABP.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293
FT CHAIN 294 467
FT PROPEP 468 474
FT CHAIN 475 >591
FT CA_BIND 19 32
FT CA_BIND 62 73
FT NON_TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

```

alignment\_scores:

Quality:	114.50	Length:	113
----------	--------	---------	-----

Ratio: 1.684 Gaps: 4  
Percent Similarity: 60.177 Percent Identity: 28.319

alignment\_block:

US-09-806-382A-1 x Q01720 ..

Align seg 1/1 to: Q01720 from: 1 to: 591

```

1 ATGTTGACCGAGCTGGAGAAAGCCCTTGAACCTATATCATGACGTCTACCA 50
  ||| ||| |||||: : : : : : : : : : : : : : : : : : : : :
1 MetSerThrLeuLeuGluAsn...IlePheAlaIleIleAsnLeuPheLys 16
51 CAAGTACTCCCTGATAAAGGGGAATTTCCATGCCCTCTACAGGGATGACC 100
  : : : : : ||| : : : : : : : : : : : : : : : : : : : :
16 sGlnTyrSerLysLysAspLysAsnThrAspThrLeuSerLysLysGluL 33
101 TGAAGAAATGCTAGACAGCCAGTGTCTCAGTATATCAGGAAAAAGGAT 150
  |||||: : : : : ||| : : : : : : : : : : : : : : : : : :
33 euLysGluLeuLeuGluLysGluPheArgGlnIleLeuLysAsnProasp 49
151 .....GCAGACGTCTGTTCAAAGAGTTGGATATCAACACTGA 188
  : : : : : : : : : : : : : : : : : : : : : : : : : :
50 AspProAspMetValAspValPheMetAspHisLeuAspIleAspHisAs 66
189 TGGTGCAGTTACTCTCCAGAGTCTCATCTGGTATGAATGATGGC. 237
  : : : : : ||| |||||: : : : : : : : : : : : : : : :
66 nLysLysIleAspPheThrGluPheLeuLeuMetValPheLysLeuAlaG 83
238 .....GTGGCAGCCCCAC 249
  : : : : : ||| : : : : : : : : : : : : : : : : : :
83 InAlaTyrTyrGluSerThrArgLysGluAsnLeuProIleSerGlnHis 99
250 .....AAAAAAGCCATGAGAAAGCCCAAAAGAG 279
  ||| |||||: : : : : : : : : : : : : : : : : : : :
100 LysHisArgLysHisSerHisHisAspLysHisGluAsp 112

```

seq\_name: sp\_human:Q9H4U2

seq\_documentation\_block:

```

ID Q9H4U2 PRELIMINARY; PRT; 687 AA.
AC Q9H4U2;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE D114N1.1.1 (PROFLAGGRIN 5' END) (FRAGMENT).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR HSSP; P02593; ICDM.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; F1laggrin.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PRO0487; F1LAGGRIN.
DR SMART; SM00054; EFh; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
FT NON_TER 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

```

alignment\_scores:

Quality: 114.50 Length: 113  
Ratio: 1.684 Gaps: 4  
Percent Similarity: 60.177 Percent Identity: 28.319

alignment\_block:

US-09-806-382A-1 x Q9H4U2 ..

Align seg 1/1 to: Q9H4U2 from: 1 to: 687

```

1 ATGTTGACCGAGCTGGAGAAAGCCCTTGAACCTATATCATGACGTCTACCA 50
  ||| ||| |||||: : : : : : : : : : : : : : : : : : : :
1 MetSerThrLeuLeuGluAsn...IlePheAlaIleIleAsnLeuPheLys 16
51 CAAGTACTCCCTGATAAAGGGGAATTTCCATGCCCTCTACAGGGATGACC 100
  : : : : : ||| : : : : : : : : : : : : : : : : : : :
16 sGlnTyrSerLysLysAspLysAsnThrAspThrLeuSerLysLysGluL 33
101 TGAAGAAATGCTAGACAGCCAGTGTCTCAGTATATCAGGAAAAAGGAT 150
  |||||: : : : : ||| : : : : : : : : : : : : : : : : : :
33 euLysGluLeuLeuGluLysGluPheArgGlnIleLeuLysAsnProasp 49
151 .....GCAGACGTCTGTTCAAAGAGTTGGATATCAACACTGA 188
  : : : : : : : : : : : : : : : : : : : : : : : : : :
50 AspProAspMetValAspValPheMetAspHisLeuAspIleAspHisAs 66
189 TGGTGCAGTTACTCTCCAGAGTCTCATCTGGTATGAATGATGGC. 237
  : : : : : ||| |||||: : : : : : : : : : : : : : : :
66 nLysLysIleAspPheThrGluPheLeuLeuMetValPheLysLeuAlaG 83
238 .....GTGGCAGCCCCAC 249
  : : : : : ||| : : : : : : : : : : : : : : : : : :
83 InAlaTyrTyrGluSerThrArgLysGluAsnLeuProIleSerGlnHis 99
250 .....AAAAAAGCCATGAGAAAGCCCAAAAGAG 279
  ||| |||||: : : : : : : : : : : : : : : : : : : :
100 LysHisArgLysHisSerHisHisAspLysHisGluAsp 112

```

OM of: US-09-806-382A-1 to: A\_Geneseq\_032802: \* out\_format : pfs  
Date: Sep 9, 2002 3:10 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=frame\_m2p.model -DEV=xlip  
-Q=/cgn2\_1/USPTO\_pool/US09806382/runat\_09092002\_143847\_15354/app\_query.fasta\_1.752  
-DB=A\_Geneseq\_032802 -QPM=fastan -SUFFIX=rag -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
-GAPOP=6.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELCP=6.000 -DELCP=7.000 -START=1 -MATRIX=blsuum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09806382 @CGNL\_1\_120 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLIPX -WAIT -THREADS=1

Search information block:

Query: US-09-806-382A-1  
Query length: 282  
Database: A\_Geneseq\_032802: \*  
Database sequences: 747574  
Database length: 111073796  
Search time (sec): 103.790000

Sequence	Strd Orig	ZScore	EScore	Len	! Documentation
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAW17061 +		485.00	1192.05	3.2e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:AAW60177 +		485.00	1192.05	3.2e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW45538 +		485.00	1192.05	3.2e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW87635 +		485.00	1192.05	3.2e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW44595 +		485.00	1192.05	3.2e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW31906 +		485.00	1192.05	3.2e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW31910 +		485.00	1192.05	3.2e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW87376 +		485.00	1192.05	3.2e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW79360 +		426.00	1038.52	6.0e-58	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:AAW50052 +		317.00	778.16	5.2e-35	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:AAW2429 +		295.00	720.33	6.4e-34	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:AAW44594 +		295.00	720.33	6.4e-34	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:AAW03563 +		160.00	384.48	3.1e-13	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:AAW03564 +		158.00	379.51	6.0e-13	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAW24137 +		158.00	379.51	6.0e-13	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW45542 +		158.00	379.51	6.0e-13	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW31907 +		158.00	379.51	6.0e-13	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW31908 +		158.00	379.51	6.0e-13	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW31911 +		158.00	379.51	6.0e-13	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW31909 +		156.00	374.65	1.1e-12	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW45544 +		154.00	369.23	2.2e-12	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW45545 +		154.00	369.23	2.2e-12	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW12007 +		154.00	367.40	2.3e-12	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW90764 +		153.00	367.32	2.9e-12	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW90765 +		153.00	367.32	2.9e-12	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:AAW01826 +		151.00	362.23	5.5e-12	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAW93819 +		145.00	346.98	3.8e-11	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW45531 +		145.00	346.98	3.8e-11	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW40258 +		145.00	346.98	3.8e-11	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW27582 +		135.00	322.01	9.2e-10	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW601409 +		134.00	322.16	1.2e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW31930 +		134.00	317.51	1.4e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAW17062 +		133.50	316.36	1.6e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:AAW60178 +		133.50	316.36	1.6e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:AAW45615 +		133.50	316.36	1.6e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW45539 +		133.50	316.36	1.6e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAW87637 +		133.50	316.36	1.6e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW44613 +		133.50	316.36	1.6e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW31905 +		133.50	316.36	1.6e-09	
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW35994 +		133.50	313.33	1.8e-09	

/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:AAW46607 + 133.00 317.49 1.7e  
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW44612 + 131.50 311.48 3.0e  
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAW6133 + 130.00 310.04 4.5e  
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAW20560 + 126.00 299.00 1.7e  
/SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:AAW80453 + 126.00 299.00 1.7e

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAW17061

seq\_documentation\_block:

ID AAW17061 standard; Protein; 93 AA.

AC AAW17061;

DT 16-JUL-1997 (first entry)

DE Human multidrug resistance protein 8 (MRP8).

KW Genetic engineering; MRP; multidrug resistance protein; transgenic;

KW animal model; cell death inhibition; apoptosis; cell proliferation;

KW HIV; human immunodeficiency virus; cancer; cystic fibrosis; neoplasia.

OS Homo sapiens.

PN US5614397-A.

PD 25-MAR-1997.

PR 22-FEB-1994; 94US-0200016.

PR 22-FEB-1994; 94US-0200016.

PA (STRD ) UNIV LEIAND STANFORD JUNIOR.

PI Lagasse E, Weissman I;

DR WPI: 1997-224943/20.

XX N-FSDB; AAT68321.

PT Increasing life-span of mammalian haemato-lymphoid cells by  
transferring stem cells - with construct contg. cell-specific  
transcription initiator and gene encoding protein that increases  
lifetime, useful for drug screening and treatment

PS Example 1; Column 27-28; 34pp: English.

XX AAW17061 is the human MRP8 protein. The transcriptional initiator of the  
MRP8 gene was used in a construct for expressing an open reading frame  
that increases the lifespan of a mammalian haematolymphoid cell, e.g. the  
mammalian bcl-2 gene, a CTR (cystic fibrosis transmembrane regulator)  
gene, the herpes virus thymidine kinase gene or an oncogene.  
XX Haematolymphoid cells are especially neutrophils and the construct  
doubles the lifespan of transgenic cells. Transgenic cells or  
transgenic animals produced are used for screening for substances and  
treatments that prevent or promote cell death. They can also be returned  
to the patient to modulate apoptosis, i.e. in the treatment of disorders  
related to abnormal cell proliferation or death. Typical applications  
are treatment of viral diseases, including HIV; cancer and cystic  
fibrosis.

XX Sequence 93 AA;

alignment\_scores: Quality: 485.00 Length: 93

Ratio: 5.215 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block: US-09-806-382A-1 x AAW17061 ..

Align seg 1/1 to: AAW17061 from: 1 to: 93

1 ANGTGACGAGCTGAGAAAGCCCTTCACTCATGACGACTCTACCA 50

```
|||||
1 MetLeuThrGluLeuGluLysAlaLeuAsnSerIleIleAspValTyrH 17
51 CAAGTACTCCCTGATAAAGGGGAATTTCCATGCCGTCTACAGGGATGACC 100
17 slYsTyrSerLeuIleIleYsGlyAsnPhenIleAlaValTyrArgAspAspL 34
101 TGAAGAAATTTGCTAGAGACCGAGTGTCCCTCAGTATATCAGGAAAAAGG 150
34 euLysLysLeuLeuGluThrGluCysProGlnTyrIleArgLysLysGly 50
151 GCAGACGTGCTGGTTCAAGAGTTGGATATCAACACTGATGTCAGCTTAA 200
51 AlaAspValTirPhelYsGluLeuAspIleAsnThrAspGlyAlaValAs 67
201 CTTCCAGAGATTCTCATTTCTGGTGATAAAGATGGCGGTGGCAGCCACA 250
67 nPhegInGluPhelLeuIleLeuValIleLysMetGlyValAlaAlaHisL 84
251 AAAAAAGCCATGAGAAAGCCACAAGAG 279
84 yslYsSerHisGluGluSerHisLysGlu 93

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:AAW60177
seq_documentation_block:
ID AAW60177 standard; Protein: 93 AA.
XX
AC AAW60177;
XX
DT 03-SEP-1998 (first entry)
XX
DE Human calprotectin subunit MRP-8 protein.
XX
KW Human; MRP-8; MRP-14; calprotectin; mineral precipitate; struvite;
KW calcium phosphate; kidney stone; renal calculi; struvite stone;
KW urinary tract infection.
XX
OS Homo sapiens.
XX
FN US5776348-A.
XX
PD 07-JUL-1998.
XX
PE 07-FEB-1995; 95US-0385241.
XX
PR 07-FEB-1995; 95US-0385241.
XX
PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Asakura H, Dretler SP, Orme-Johnson WH, Selengut JD;
XX
DR WPI: 1998-397914/34.
XX
DR N-PSDB: AAV34697.
XX
PT Inhibiting kidney stone formation - uses the protein calprotectin
XX
PS Example 1; Columns 19-20; 19pp: English.
XX
XX This represents a human calprotectin subunit MRP-8. This is used in
CC a method for inhibiting the formation of a mineral precipitate in a
CC solution which comprises providing a solution comprising component ions
CC of the mineral precipitate or its precursors and contacting the solution
CC with an effective amount of isolated calprotectin, or a derivative of it.
CC The method is useful for the inhibition of kidney stone formation (renal
CC calculi). Kidney stones are concentrations of inorganic and organic salts
CC that develop through crystal nucleation, aggregation and growth in the
CC kidneys, which then can block the ureter and if not passed to the
CC bladder, grow and become symptomatic. Struvite stones (MgNH4PO4) are
CC commonly found after urinary tract infection. The mechanisms of this and
CC other stone formation is unclear, but some urine compositions are known
CC to inhibit formation. One of the major components of these is a protein
CC factor calprotectin, which inhibits mineral precipitation. Calprotectin
```

```
CC can also be used to raise antibodies, which in turn can be used to detect
CC the protein in samples. The levels of calprotectin found in samples can
CC be compared to levels found in normal humans, and thus assuming increased
CC calprotectin correlates to indication of kidney stone formation, this
CC procedure can be used as a diagnostic tool.
XX
SO Sequence 93 AA;

alignment_scores:
Quality: 485.00 Length: 93
Ratio: 5.215 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-806-382a-1 x AAW60177 ..
Align seg 1/1 to: AAW60177 from: 1 to: 93

1 ATGTGACCGAGCGTGAAGAGCCCTGAACCTATATCATCGAGCGTACCA 50
|||||
1 MetLeuThrGluLeuGluLysAlaLeuAsnSerIleIleAspValTyrH 17
17 slYsTyrSerLeuIleIleYsGlyAsnPhenIleAlaValTyrArgAspAspL 34
51 CAAGTACTCCCTGATAAAGGGGAATTTCCATGCCGTCTACAGGGATGACC 100
|||||
101 TGAAGAAATTTGCTAGAGACCGAGTGTCCCTCAGTATATCAGGAAAAAGG 150
34 euLysLysLeuLeuGluThrGluCysProGlnTyrIleArgLysLysGly 50
151 GCAGACGTGCTGGTTCAAGAGTTGGATATCAACACTGATGTCAGCTTAA 200
51 AlaAspValTirPhelYsGluLeuAspIleAsnThrAspGlyAlaValAs 67
201 CTTCCAGAGATTCTCATTTCTGGTGATAAAGATGGCGGTGGCAGCCACA 250
67 nPhegInGluPhelLeuIleLeuValIleLysMetGlyValAlaAlaHisL 84
251 AAAAAAGCCATGAGAAAGCCACAAGAG 279
84 yslYsSerHisGluGluSerHisLysGlu 93

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:AA845538
seq_documentation_block:
ID AA845538 standard; Protein: 93 AA.
XX
AC AA845538;
XX
DT 22-FEB-2001 (first entry)
XX
DE Human S100A8 protein.
XX
KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
KW calcium-binding protein; calcium homeostasis; cardiac muscle;
KW pumping capacity; myocardial cell; systolic calcium ion release;
KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
KW valve defect.
XX
OS Homo sapiens.
XX
FN DE19915485-A1.
XX
PD 19-OCT-2000.
XX
PE 07-APR-1999; 99DE-1015485.
XX
PR 07-APR-1999; 99DE-1015485.
XX
PA (KATU/) KATUS H A.
XX
PA (REMP/) REMPIIS A.
XX
```

PI Katus HA, Remppis A;  
XX  
XX MPI: 2000-673510/66.  
DR N-PSDB; AAC81808.  
XX

PT Composition containing S100 protein, corresponding nucleic acid or  
PT vector, useful for treating cardiomyopathy and cardiac insufficiency  
XX  
XX Claim 35; Page 15; 36pp; German.

XX This invention describes a novel composition for treating primary or  
XX secondary cardiomyopathy or cardiac insufficiency contains at least one  
XX S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or  
XX fragments, or a gene transfer vector containing (II), optionally  
XX formulated with auxiliaries and/or carriers. (I) are calcium-binding  
XX proteins involved in calcium homeostasis, so their overexpression in  
XX cardiac muscle will improve pumping capacity (and overall capacity) of  
XX the heart. In cultured myocardial cells they increase the contraction and  
XX relaxation rates associated with increased systolic calcium ion release  
XX from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are  
XX used to treat cardiomyopathy (CMP) where inherited or caused by  
XX spontaneous mutations and ischemic CMP caused by arteriosclerosis,  
XX dilative CMP caused by toxic/infectious disease, cardiac disease caused  
XX by pulmonary and/or arterial hypertension, and structural disease caused  
XX by rhythm disorders or valve defects, generally any condition associated  
XX with reduced contractile force. Unlike calmodulin, which is expressed  
XX ubiquitously, (I) show tissue-specific expression and treat the  
XX underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac  
XX disease.

XX Sequence 93 AA;  
SQ

alignment\_scores: Quality: 485.00 Length: 93  
Ratio: 5.215 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-806-382a-1 x AAB45538 ..

Align seg 1/1 to: AAB45538 from: 1 to: 93

1 ATGTTGACCGAGCTGGAGAAAGCCTGAACCTATCATCGAGCTTACCA 50  
|||||  
1 MetLeuThrGluLeuGluAlaLeuAsnSerIleLeuValTyrHn 17  
51 CAAGTACTCCCTGATAAAGGGGAATTTCACATGCCCTACAGGGATGACC 100  
|||||  
17 slsTYSerLeuIleLysGluAsnPhenAlaValTyrArgAspAspL 34  
101 TGAAGAAATGCTAGAGACCGAGTGTCTCAGTATTCAGGAAAGGCT 150  
|||||  
34 euLysLysLeuLeuGluThrGluCysProGlnTyrIleArgLysGly 50  
151 GCAGAGCTGTGTTCAAGAGTTGATATCAACATGATGCTGCTACGTTAA 200  
|||||  
51 AlaAspValIlePheLysGluLeuAspIleAsnThrAspGlyAlaValAs 67  
201 CTTCACGAGAGTTCTCATCTGTGTTGATTAAGATGGCGTGGCAGCCACA 250  
|||||  
67 nPheGlnGluPheLeuIleLeuValIleLysMetGlyValAlaIleHnL 84

251 AAAAAAGCCATGAAGAAAGCCACAAAGAG 279

84 yLysSerHisGluGluSerHisLysGlu 93

seq\_name: /SIDS1/ycgdata/hold-geneseq/genesegp-emb1/AA2000.DAT: AAY87636

seq\_documentation\_block:

ID AAY87636 standard; Protein: 93 AA.

XX  
XX AAY87636;

XX 04-AUG-2000 (first entry)  
DT  
XX  
DE Human calcium-binding protein #1.  
XX

XX Calcium-binding protein; granule release; calgranulin; human;  
XX vascular membrane growth; adult respiratory distress syndrome;  
XX acute myocardial infarction; ischemic reperfusion disorder;  
XX glomerulonephritis; rheumatoid arthritis; chronic bronchitis;  
XX cerebral vascular disorder; asthma; peripheral circulation disturbance;  
XX angina pectoris; hypertension; multiple sclerosis.  
XX

XX Homo sapiens.  
XX  
XX W0200018970-A1.  
XX  
XX 06-APR-2000.  
XX

XX 28-SEP-1999; 99WO-JP05302.  
XX

XX 29-SEP-1998; 98JP-0274574.  
XX

XX (ASAH ) ASAH KASEI KOGYO KK.  
XX

XX Seto M, Fukuda K;  
XX

XX MPI: 2000-293189/25.  
XX

XX N-PSDB; AAA11969.  
XX

PT Controlling the release of granules from cell system using activated  
PT calgranulin for screening substances for granule activating or  
PT inhibiting activity  
XX  
XX Disclosure; Page 37-38; 42pp; Japanese.

XX This invention describes a novel method for controlling the release of  
XX granules by treating a cell system to increase or decrease activated  
XX calgranulin to enhance or depress the release of granules. Calgranulin  
XX is a calcium binding protein and can be used for controlling the release  
XX of granules from a cell system e.g. those involved with the inhibition of  
XX vascular membrane growth. Vascular membrane growth is associated with  
XX e.g. adult respiratory distress syndrome, acute myocardial infarction due  
XX to ischemic reperfusion disorders, glomerulonephritis, rheumatoid  
XX arthritis, chronic bronchitis, cerebral vascular disorders, asthma,  
XX peripheral circulation disturbance, angina pectoris, hypertension and  
XX multiple sclerosis. The new method is used for screening substances for  
XX their ability to activate or inhibit the release of granules. This  
XX sequence represents a human calcium-binding protein which is described in  
XX the method of the invention.

SQ Sequence 93 AA;

alignment\_scores: Quality: 485.00 Length: 93  
Ratio: 5.215 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-806-382a-1 x AAY87636 ..

Align seg 1/1 to: AAY87636 from: 1 to: 93

1 ATGTTGACCGAGCTGGAGAAAGCCTGAACCTATCATCGAGCTTACCA 50  
|||||  
1 MetLeuThrGluLeuGluAlaLeuAsnSerIleLeuValTyrHn 17  
51 CAAGTACTCCCTGATAAAGGGGAATTTCACATGCCCTACAGGGATGACC 100  
|||||  
17 slsTYSerLeuIleLysGluAsnPhenAlaValTyrArgAspAspL 34  
101 TGAAGAAATGCTAGAGACCGAGTGTCTCAGTATTCAGGAAAGGCT 150  
|||||

34 euLysLeuLeuGluThrGluCysProGlnTyrIleArgLysLysGly 50  
151 GCAGACGTCGTGTTCAAGAGTTGGATATCAACACTGATGTGCAGTTAA 200  
51 AlaAspValTrpPheLysGluLeuAspIleAsnThrAspGlyAlaValAs 67  
201 CTTCCAGAGAGTTCCTCATTTCTGTGATTAAGATGGCGCGCCAGCCACA 250  
67 nPheGlnGluPheLeuIleLeuValIleLysMetGlyValAlaAlaHisL 84  
251 AAAAAGCCATGAGAAAGCCACAAAGAG 279  
84 yslYSerHisGluGlnSerHisLysGln 93

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:ABB44595

seq\_documentation\_block:

ID ABB44595 standard; Protein: 93 AA.  
AC ABB44595;  
XX  
XX 25-JAN-2002 (first entry)  
DT  
XX Human wound healing related polypeptide SEQ ID NO 54.  
DE  
XX Human: mouse; vulnary; dermatological; skin disorder; wound healing;  
KW gene therapy.  
XX  
XX Homo sapiens.  
OS  
XX CA2325226-A1.  
PN  
XX 17-MAY-2001.  
PD  
XX 16-NOV-2000; 2000CA-2325226.  
PE  
XX 17-NOV-1999; 99DE-1055349.  
PR 17-DEC-1999; 99US-0172511.  
PR 20-JUN-2000; 2000DE-1030149.  
XX  
XX (SWIT-) SWITCH BIOTECH AG.  
PA  
XX  
XX Regendogen J, Wolf E, Goppelt A, Werner S, Halle J;  
PI  
XX WPI: 2001-433142/47.  
DR  
XX  
XX Use of novel polypeptide or its variant or nucleic acid encoding the  
PT polypeptide for diagnosing and/or preventing and/or treating skin  
PT disorders and/or treatment in wound healing or for identifying active  
PT substances -  
XX  
XX  
PS Claim 3; Page 220; 265pp; English.  
XX  
XX The invention relates to the use of a polypeptide (ABB44544-ABB44601,  
CC ABB44606-ABB44623) or its variant or encoding nucleic acid  
CC (ABA81990-ABA81995, ABA82016-ABA82032) with vulnary and/or  
CC dermatological activity for the diagnosis, prevention and treatment of  
CC skin disorders and treatment in wound healing or for the identification  
CC of pharmacologically active substances. The nucleic acids are useful in  
CC gene therapy.  
CC Note: The printed sequence listing for this specification was incomplete,  
CC terminating part way through SEQ ID NO 106. The remaining data was  
CC obtained from EPO data for an equivalent patent (EP1114862).  
XX  
XX Sequence 93 AA;  
SQ

alignment\_scores:                      length: 93  
                    Quality: 485.00  
                    Ratio: 5.215  
Percent Similarity: 100.000            Percent Identity: 100.000  
alignment\_block:  
:

US-09-806-382a-1 x ABB44595 ..

Align seg 1/1 to: ABB44595 from: 1 to: 93

1 ATGTTGACCCAGACGTGGAGAAAGCTTGAACCTGATCATGAGCTTACCA 50  
1 MetLeuThrGluLeuGluLysAlaLeuAsnSerIleIleAspValTyrn 17  
51 CAACTACTCCCTGATTAAGGGAATTTCCATGCGCGCTACAGGAGAGACC 100  
17 slYsTyrSerLeuIleLysGlyAsnPheHisAlaValTyrArgAspAspL 34  
101 TGAAGAAATTCGTAGAGCCAGTGTCTCAGTATATCAAGAAAGGCT 150  
34 euLysLysLeuLeuGluThrGluCysProGlnTyrIleArgLysLysGly 50  
151 GCAGACGTCGTGTTCAAGAGTTGGATATCAACACTGATGTGCAGTTAA 200  
51 AlaAspValTrpPheLysGluLeuAspIleAsnThrAspGlyAlaValAs 67  
201 CTTCCAGAGAGTTCCTCATTTCTGTGATTAAGATGGCGCGCCAGCCACA 250  
67 nPheGlnGluPheLeuIleLeuValIleLysMetGlyValAlaAlaHisL 84  
251 AAAAAGCCATGAGAAAGCCACAAAGAG 279  
84 yslYSerHisGluGlnSerHisLysGln 93

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAB31906

seq\_documentation\_block:

ID AAB31906 standard; Protein: 93 AA.  
XX  
XX AAB31906;  
XX  
XX 15-MAY-2001 (first entry)  
DT  
XX Amino acid sequence of a human protein.  
DE  
XX  
XX Human: perlecan; retinol-binding plasma protein; calgranulin B; vaccine;  
KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;  
KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;  
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
KW rheumatoid polyarthritis; lupus erythematosus; gene therapy.  
XX  
XX Homo sapiens.  
OS  
XX WO200105422-A2.  
PN  
XX 25-JAN-2001.  
PD  
XX 17-JUL-2000; 2000WO-FR02057.  
PE  
XX 15-JUL-1999; 99FR-0009372.  
PR  
XX (INNER) BIOMERIEUX STELHYS.  
PA  
XX Roeklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;  
PI WPI: 2001-159475/16.  
DR  
XX  
XX Detecting, preventing and treating degenerative, neurological and  
PT autoimmune diseases, particularly multiple sclerosis, using specified  
PT polypeptides or related nucleic acid or ligand -  
XX  
XX  
PS Claim 1; Page 166; 209pp; French.  
XX

The present sequence represents a human protein, which is used in the  
method of the invention. The specification describes a method which uses  
at least one polypeptide or polynucleotide sequence belonging to the  
perlecan, precursor of the retinol-binding plasma protein, precursor of  
the ganglioside GM2 activator, calgranulin B or saposin B protein  
families. The method is used for detecting, preventing or treating a

CC degenerative, neurological and/or auto-immune disease. The  
CC polynucleotides and polypeptides are used for diagnosis, prognosis,  
CC prevention and treatment of multiple sclerosis (in its various forms  
CC and phases). They may also be useful in cases of e.g. Alzheimer's and  
CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid  
CC polyarthritis and lupus erythematosus, including use as vaccines and  
CC in gene therapy (expression of sense or antisense sequences). They can  
CC also be used to assess efficacy of potential therapeutic agents,  
CC particularly compounds that reduce or inhibit toxicity towards glial  
CC cells.

SO Sequence 93 AA:

alignment\_scores:                      Length: 93  
                    Quality: 485.00                      Gaps: 0  
                    Ratio: 5.215                      Percent Identity: 100.000  
Percent Similarity: 100.000                      Percent Identity: 100.000

alignment\_block:

US-09-806-382a-1 x AAB31906 ..

Align seg 1/1 to: AAB31906 from: 1 to: 93

```
1 ATGTTGACGAGCTGGAGAAAGCCTGAACCTATCATGACGCTTACCA 50
|||||
1 MetLeuThrGluLeuGluAlaLeuAsnSerIleIleAspValTyrH 17
51 CAAGTACTCCCTGATTAAGGGGAATTTCCATGCCGTCTACAGGATGACC 100
|||||
17 sLysTyrSerLeuIleLysGluAsnPhenAlaValTyrArgAspAsp 34
101 TGAAGAAATTCGTAAGACCGAGCTGCTCAGTATTCAGGAAAAAGGT 150
|||||
34 euLysLysLeuLeuGluThrGluCysProGlnTyrIleArgLysGly 50
151 GCAGACGCTGCTGTTCAAGAGTGGATATCAACATGATGGTGCAGCTTAA 200
|||||
51 AlaAspValTyrPheLysGluLeuAspIleAsnThrAspGlyAlaValAs 67
201 CTTCAGAGAGTTCCTCATTCGTGTGATTAAGATGGCGTGGCAGCCGACA 250
|||||
67 nPheGlnGluPheLeuIleLeuValIleLysMetGlyValAlaIleAsn 84
251 AAAAAAGCCATGAAGAAAGCCACAAAGAG 279
|||||
84 yLysSerHisGluGluSerHisLysGlu 93
```

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAB31910

seq\_documentation\_block:

ID AAB31910 standard; Protein: 93 AA.

AC AAB31910;

DT 15-MAY-2001 (first entry)

DE Amino acid sequence of a human protein.

XX Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;  
XX ganglioside GM2 activator; saposin B; degenerative disease; glial cell;  
XX neurological disease; auto-immune disease; multiple sclerosis; toxicity;  
XX Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
XX rheumatoid polyarthritis; lupus erythematosus; gene therapy.

OS Homo sapiens.

PN WO200105422-A2.

XX 25-JAN-2001.

PF 17-JUL-2000; 2000WO-FR02057.

XX

PR 15-JUL-1999; 99FR-0009372.  
XX  
XX (INMR ) BIOMERIEUX STELHYS.

PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;  
XX WPI; 2001-159475/16.

DR Detecting, preventing and treating degenerative, neurological and  
XX autoimmune diseases, particularly multiple sclerosis, using specified  
XX polypeptides or related nucleic acid or ligand

PS Claim 1; Page 168; 209pp; French.

XX The present sequence represents a human protein, which is used in the  
XX method of the invention. The specification describes a method which uses  
XX at least one polypeptide or polynucleotide sequence belonging to the  
XX perlecan, precursor of the retinol-binding plasma protein, precursor of  
XX the ganglioside GM2 activator, calgranulin B or saposin B protein  
XX families. The method is used for detecting, preventing or treating a  
XX degenerative, neurological and/or auto-immune disease. The  
XX polynucleotides and polypeptides are used for diagnosis, prognosis,  
XX prevention and treatment of multiple sclerosis (in its various forms  
XX and phases). They may also be useful in cases of e.g. Alzheimer's and  
XX Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid  
XX polyarthritis and lupus erythematosus, including use as vaccines and  
XX in gene therapy (expression of sense or antisense sequences). They can  
XX also be used to assess efficacy of potential therapeutic agents,  
XX particularly compounds that reduce or inhibit toxicity towards glial  
XX cells.

SO Sequence 93 AA:

alignment\_scores:                      Length: 93  
                    Quality: 485.00                      Gaps: 0  
                    Ratio: 5.215                      Percent Identity: 100.000  
Percent Similarity: 100.000                      Percent Identity: 100.000

alignment\_block:

US-09-806-382a-1 x AAB31910 ..

Align seg 1/1 to: AAB31910 from: 1 to: 93

```
1 ATGTTGACGAGCTGGAGAAAGCCTGAACCTATCATGACGCTTACCA 50
|||||
1 MetLeuThrGluLeuGluAlaLeuAsnSerIleIleAspValTyrH 17
51 CAAGTACTCCCTGATTAAGGGGAATTTCCATGCCGTCTACAGGATGACC 100
|||||
17 sLysTyrSerLeuIleLysGluAsnPhenAlaValTyrArgAspAsp 34
101 TGAAGAAATTCGTAAGACCGAGCTGCTCAGTATTCAGGAAAAAGGT 150
|||||
34 euLysLysLeuLeuGluThrGluCysProGlnTyrIleArgLysGly 50
151 GCAGACGCTGCTGTTCAAGAGTGGATATCAACATGATGGTGCAGCTTAA 200
|||||
51 AlaAspValTyrPheLysGluLeuAspIleAsnThrAspGlyAlaValAs 67
201 CTTCAGAGAGTTCCTCATTCGTGTGATTAAGATGGCGTGGCAGCCGACA 250
|||||
67 nPheGlnGluPheLeuIleLeuValIleLysMetGlyValAlaIleAsn 84
251 AAAAAAGCCATGAAGAAAGCCACAAAGAG 279
|||||
84 yLysSerHisGluGluSerHisLysGlu 93
```

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAM78376

seq\_documentation\_block:

ID AAM78376 standard; Protein: 171 AA.

XX

```
AC AAM78376;
XX
XX 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1038.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
XX MO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX 20-JUN-2000; 2000US-0598075.
XX 19-JUL-2000; 2000US-0620325.
XX 01-SEP-2000; 2000US-0654936.
XX 15-SEP-2000; 2000US-0663561.
XX 20-OCT-2000; 2000US-0693325.
XX 30-NOV-2000; 2000US-0728422.
XX
XX (HXSE-) HXSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
XX N-PSDB; AAK51509.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
XX Claim 20; Page 3275-3276; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 171 AA;

alignment_scores:
    Quality: 485.00      Length: 93
    Ratio: 5.215        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-806-382A-1 x AAM78376 ..

Align seg 1/1 to: AAM78376 from: 1 to: 171

1 AMGTTCAGCCGAGCTGGAGAAAGCCTTGAACCTCATCATGAGCTCTTACCA 50
|||||
79 MetLeuThrGluLeuGlnIuysAlaLeuAsnSerIleIleAspValTyrH 95
```

```
51 CAAGTACTCCCTGATAAAGGGAATTTCCATGCCGCTCTACAGGGATGACC 100
|||||
95 slystYrserIleuIleIySGIyAsnPhetSAlaValIYrAtgAspAspL 112
101 TGAAGAAATTCCTAGAGACCGAGTCTCTCAGTATATCAAGAAAAAGGT 150
|||||
112 euIysIysLeuLeuGluTnHrGluCySProGInTyrIleatGlySgIy 128
151 GCAGAGCTCTGGTTCAAGAGTTGGATATCAACACGTGATGGTCCAGTTAA 200
|||||
129 AlaAspValTrrPheIySGIuLeuAspIleAsnTrrAspIyAlaValAs 145
201 CTTCAGAGATTCCATTCCTGCTGATAAAGATGGCGTGGCAGCCACAGA 250
|||||
145 nPhetGInGluPheLeuIleIeuValIleIySMeTGIyValAlaIaIeHIsL 162
251 AAAAAAGCCATGAAGAACCCACAAAGAG 279
|||||
162 yIysSerHIsGInGluSerHIsIySGIu 171

seq_name: /SIDSL/gcdata/hold-geneseq/geneseqp-emb1/AA2001.DAT: AAM79360
seq_documentation_block:
ID AAM79360 standard; Protein; 179 AA.
XX
XX AAM79360;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human protein SEQ ID NO 3006.
XX
XX Human protein SEQ ID NO 3006.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
XX Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX 20-JUN-2000; 2000US-0598075.
XX 19-JUL-2000; 2000US-0620325.
XX 01-SEP-2000; 2000US-0654936.
XX 15-SEP-2000; 2000US-0663561.
XX 20-OCT-2000; 2000US-0693325.
XX 30-NOV-2000; 2000US-0728422.
XX
XX (HXSE-) HXSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
XX N-PSDB; AAK52493.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
XX Claim 229; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
```



peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.  
 Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAK80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

Sequence 179 AA:

alignment\_scores:  
 Quality: 426.00 Length: 92  
 Ratio: 4.897 Gaps: 0  
 Percent Similarity: 94.365 Percent Identity: 91.304

alignment\_block:  
 US-09-806-382a-1 x AAM79360 ..

Align seg 1/1 to: AAM79360 from: 1 to: 179

```

1  AAGTTGACCGAGCTGGAGAAAGCCTTGAACCTATCATCGACGCTACCA 50
  ||||||||||||||||||||||||||||||||||||||||||||||||
15 MetLeuThrGluLeuGluLysAlaLeuAsnSerIleIleAspValTyrH 31
51 CAAGTACCTCCGATTAAGGGGAATTCATGCCGCTCAGAGGATGACC 100
  ||||||||||||||||||||||||||||||||||||||||||||||||
31 sLysTyrSerIleuIleLysGlyAsnPhenIleAlaValTyrArgAspSpl 48
101 TGAAGAAATTTGCTAGACCGAGTCTCTCAGTATATCAGAAAAAGGCT 150
  ||||||||||||||||||||||||||||||||||||||||||||||||
48 euLysLysLeuLeuGluThrGluCysProGlnTyrIleArgLysLysGly 64
151 GCAGACGCTGGTTCAAGAGTTGGATATCACACTGATGGTGCAGTTAA 200
  ||||||||||||||||||||||||||||||||||||||||||||||||
65 AlaAspValTyrPheLysGluLeuAspIleAsnThrAspGlyAlaValAs 81
201 CTTCACGAGATTCCTCATCTGCTGATTAAGATGGCGTGGACCCACCA 250
  ||||||||||||||||||||||||||||||||||||||||||||||||
81 nPheGlnGluPheLeuIleLeuValIleLysMetGlyValAlaIleLeuA 98
251 AAAAAAGCCATGAAGAAAGCCACAAA 276
  ::::: |||||
98 snSerIleIleAspValTyrHisLys 106

```

seq\_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT: AAP50052

seq\_documentation\_block:

ID AAP50052 standard; Protein: 66 AA.

AC AAP50052;

DT 09-SEP-1991 (first entry)

DE Sequence of human macrophage migration inhibiting factor (MIF).

KW Autoimmune disease therapy; tumour therapy; leprosy; tuberculosis.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 42 /label= S,C

XX EPI62812-A.

XX 27-OCT-1985.

XX 15-MAY-1985; 85EP-0810232.

XX 05-FEB-1985; 85CH-0000512.

PR 24-MAY-1984; 84CH-0002557.  
 PR 07-AUG-1984; 84CH-0003786.  
 PR 14-NOV-1984; 84CH-0005446.  
 XX  
 XX (CIBA ) CIBA GEIGY AG.

XX Sorg C, Burmeister G, Tarcsay L, Wiesendanger W;  
 PI WPI; 1985-298143/48.

PT Pure human macrophage migration inhibiting factor - specific  
 PT monoclonal antibodies and hybridoma cell lines, useful e.g. for  
 PT treating auto-immune disease

PS Claim 5; Page 38; 48pp; German.

CC MIF is used for increasing resistance to infection (e.g.  
 CC tuberculosis, leprosy, leishmaniasis or candidiasis) and to tumours  
 CC (esp. metastases).

SO Sequence 66 AA;

alignment\_scores:  
 Quality: 317.00 Length: 66  
 Ratio: 5.032 Gaps: 0  
 Percent Similarity: 95.455 Percent Identity: 93.939

alignment\_block:  
 US-09-806-382a-1 x AAP50052 ..

Align seg 1/1 to: AAP50052 from: 1 to: 66

```

1  AAGTTGACCGAGCTGGAGAAAGCCTTGAACCTATCATCGACGCTACCA 50
  ||||||||||||||||||||||||||||||||||||||||||||||||
1 MetLeuThrGluLeuGluLysAlaLeuAsnSerIleIleAspValTyrH 17
51 CAAGTACCTCCGATTAAGGGGAATTCATGCCGCTCAGAGGATGACC 100
  ||||||||||||||||||||||||||||||||||||||||||||||||
17 sLysTyrSerIleuIleLysGlyAsnPhenIleAlaValTyrArgAspSpl 34
101 TGAAGAAATTTGCTAGACCGAGTCTCTCAGTATATCAGAAAAAGGCT 150
  ||||||||||||||||||||||||||||||||||||||||||||||||
34 euLysLysLeuLeuGluThrGlu***ProGlnTyrIleArgLysLysGly 50
151 GCAGACGCTGGTTCAAGAGTTGGATATCAACACTGATGTCAGTT 198
  ||||||||||||||||||||||||||||||||||||||||||||||||
51 AlaAspValTyrPheLysGluLeuAspIleAsn*****AlaVal 66

```

seq\_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT: AAR22429

seq\_documentation\_block:

ID AAR22429 standard; Protein: 89 AA.

AC AAR22429;

DT 12-AUG-1992 (first entry)

DE Murine CP-10.

KW Chemotactic protein; neutrophil; monocyte; macrophage;  
 KW inflammation; autoimmune disease; delayed hypersensitivity;  
 KW CF; cystic fibrosis; emphysema; diagnosis.

XX Mus musculus.

XX WO9204376-A.

XX 19-MAR-1992.

XX 05-SEP-1991; 91WO-AU00410.

XX 05-FEB-1991; 91AU-0004463.

PR 05-SEP-1990; 90AU-0002127.  
XX  
XX (HEAR-) HEART RES INST LTD.  
XX  
PI Geczy C, Simpson RJ, Lackmann M;  
XX  
XX WPI; 1992-114301/14.  
DR N-PSDB; AAO23343.  
XX  
XX New chemotactic protein CP-10 for altering inflammatory capacity  
PT - controls inflammatory conditions, cystic fibrosis and  
PT emphysema, antagonists for detecting CP-10  
XX  
XX Claim 3; Page 36; 56pp; English.  
XX  
XX The MCP10 cDNA was derived from a Sepharose-Concanavalin A-activated  
CC spleen cell library prep. using cells from A/J mice. CP-10 has an  
CC apparent mol. wt. of ca. 10 kD and is chemotactic for neutrophils,  
CC monocytes/macrophages and/or other mammalian cells. CP-10 is useful  
CC (e.g. when admin. locally around a tumour or site of infection) to  
CC attract neutrophils etc., esp. in immunocompromised patients or  
CC patients allergic to skin test antigens. Antibodies to CP-10 are  
CC useful for control of inflammation and/or tissue damage associated  
CC with inflammation, (e.g. due to autoimmune disease or delayed  
CC hypersensitivity), cystic fibrosis or emphysema. Ab can also be used  
CC purely CP-10 and in diagnosis.  
XX  
XX Sequence 89 AA;  
SQ  
  
alignment\_scores:  
Quality: 295.00 Length: 89  
Ratio: 3.734 Gaps: 0  
Percent Similarity: 88.764 Percent Identity: 58.427  
  
alignment\_block:  
US-09-806-382A-1 x AAR22429 ..  
  
Align seg 1/1 to: AAR22429 from: 1 to: 89  
  
1 ATGTTGACCGAGCTGGAGAAAGCCTTGACCTGATCATCGACGCTACCA 50  
1 MetProserGluLeuGluLysAlaLeuSerAsnLeuIleAspValTyrHl 17  
  
51 CAAGTACTCCCTGATTAAGGGGAATTTCATGCCGCTACAGGAGTACC 100  
17 sAsnTyrSerAsnIleGlnGlyAsnHisAlaLeuTyrLysAsnSpr 34  
  
101 TGAAGAAATTCCTAGAGACCGAGTCTCCTCAGTATATCAGAAAAAGGT 150  
34 heLysLysMetValThrThrGluCysProGlnPheValGlnAsnIleAsn 50  
  
151 GCAGACGCTGGTTCAAAGAGTTGGATATCAACACTGATGGTCACTTAA 200  
51 IleGluAsnLeuPheArgGluLeuAspIleAsnSerAspAsnAlaIleAs 67  
  
201 CTTCAGAGATTCCTCATTTCTGATGAATAAGATGGCGTGGCAGCCACA 250  
67 nPheGluGluPheLeuAlaMetValIleLysValGlyValAlaSerHisL 84  
  
251 AAAAAACCATGAAGAA 267  
84 yAspSerHisLysGlu 89  
  
seq\_name: /STD1/9cgsdata/hold-geneseq/geneseq-emb1/AA2001.DAT: ABB44594  
  
seq\_documentation\_block:  
ID ABB44594 standard; Protein: 89 AA.  
XX  
XX ABB44594;  
AC  
XX  
XX 25-JAN-2002 (first entry)  
PT  
XX

DE Mouse wound healing related polypeptide SEQ ID NO 53.  
XX  
XX Human; mouse; vulnery; dermatological; skin disorder; wound healing;  
KW gene therapy.  
KW  
XX  
XX Mus musculus.  
OS  
XX  
XX CA2325226-A1.  
PN  
XX  
XX 17-MAY-2001.  
PD  
XX  
XX 16-NOV-2000; 2000CA-2325226.  
PF  
XX  
XX 17-NOV-1999; 99DE-1055349.  
PR 17-DEC-1999; 99US-0172511.  
PR 20-JUN-2000; 2000DE-1030149.  
XX  
XX (SWIT-) SWITCH BIOTECH AG.  
XX  
XX Regenbogen J, Wolf E, Goppelt A, Werner S, Halle J;  
XX  
XX WPI; 2001-433142/47.  
XX  
XX use of novel polypeptide or its variant or nucleic acid encoding the  
PT polypeptide for diagnosing and/or preventing and/or treating skin  
PT disorders and/or treatment in wound healing or for identifying active  
PT substances  
XX  
XX  
XX Disclosure: Page 219-220; 265pp; English.  
XX  
XX The invention relates to the use of a polypeptide (ABB44544-ABB44601,  
CC ABB44606-ABB44623) or its variant or encoding nucleic acid  
CC (AAB41990-ABA81995, ABA82016-ABA82032) with vulnery and/or  
CC dermatological activity for the diagnosis, prevention and treatment of  
CC skin disorders and treatment in wound healing or for the identification  
CC of pharmacologically active substances. The nucleic acids are useful in  
CC gene therapy.  
CC Note: The printed sequence listing for this specification was incomplete,  
CC terminating part way through SEQ ID NO 106. The remaining data was  
CC obtained from EPO data for an equivalent patent (EP1114862).  
XX  
XX Sequence 89 AA;  
SQ  
  
alignment\_scores:  
Quality: 295.00 Length: 89  
Ratio: 3.734 Gaps: 0  
Percent Similarity: 88.764 Percent Identity: 58.427  
  
alignment\_block:  
US-09-806-382A-1 x ABB44594 ..  
  
Align seg 1/1 to: ABB44594 from: 1 to: 89  
  
1 ATGTTGACCGAGCTGGAGAAAGCCTTGACCTGATCATCGACGCTACCA 50  
1 MetProserGluLeuGluLysAlaLeuSerAsnLeuIleAspValTyrHl 17  
  
51 CAAGTACTCCCTGATTAAGGGGAATTTCATGCCGCTACAGGAGTACC 100  
17 sAsnTyrSerAsnIleGlnGlyAsnHisAlaLeuTyrLysAsnSpr 34  
  
101 TGAAGAAATTCCTAGAGACCGAGTCTCCTCAGTATATCAGAAAAAGGT 150  
51 IleGluAsnLeuPheArgGluLeuAspIleAsnSerAspAsnAlaIleAs 67  
  
151 GCAGACGCTGGTTCAAAGAGTTGGATATCAACACTGATGGTCACTTAA 200  
51 IleGluAsnLeuPheArgGluLeuAspIleAsnSerAspAsnAlaIleAs 67  
  
201 CTTCAGAGATTCCTCATTTCTGATGAATAAGATGGCGTGGCAGCCACA 250  
67 nPheGluGluPheLeuAlaMetValIleLysValGlyValAlaSerHisL 84

251 AAAAAAGCATGAGAA 267  
|||  
84 ysaspserrhltslysglu 89

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:AAW03563

seq\_documentation\_block:

ID AAW03563 standard; Protein; 92 AA.

AAW03563;

01-MAY-1997 (first entry)

Calcium binding protein CAAFL.

Calcium binding protein; bovine; amniotic fluid; S100 protein family;  
intracellular signal transduction; squamous epithelial cell; neutrophil;  
macrophage; cancer; cancerous lesion; inflammation; neoplasia; cervix;  
squamous cell carcinoma; skin; oesophagus; CAAFL; lung; blood disease.

Bos taurus.

EP731166-A2.

11-SEP-1996.

04-DEC-1995; 95EP-0119045.

06-MAR-1995; 95JP-0070468.

06-MAR-1995; 95JP-0045564.

(HITO/) HITOMI J.

(TOFU) TONEN CORP.

Hitomi J, Kimura T, Yamaguchi K, Yamamura T;

WPI; 1996-403989/41.

N-PSDB; AAT39345.

New human or bovine calcium binding protein and related nucleic acid  
- is a marker for inflammation, neoplasia, skin and blood diseases

Claim 1; Page 21; 36pp; English.

This sequence represents the CAAFL calcium-binding protein isolated from  
bovine amniotic fluid. CAAFL belongs to the S100 protein family, which  
includes calyculin, MRP8, and MRP14. Intracellular calcium ion  
concentration is one of the key factors for intracellular signal  
transduction. The calcium signals are transduced by various  
calcium-binding proteins, such as the protein encoded by this sequence.  
CAAFL is normally expressed in squamous epithelial cells, neutrophils and  
macrophages, but atypical epithelial cells are negative for CAAFL and  
overexpression is observed in several types of cancer cells and  
neutrophils/macrophages infiltrating cancerous lesions. Detection of  
CAAFL (using antibodies in usual immunoassays) can be used to diagnose  
(or monitor) inflammation, neoplasia (particularly squamous cell  
carcinoma of the skin, oesophagus, lung and cervix), and skin and blood  
diseases.

Sequence 92 AA;

alignment\_scores:

Quality: 160.00 Length: 92  
Ratio: 2.500 Gaps: 1  
Percent Similarity: 69.565 Percent Identity: 35.870

alignment\_block:

US-09-806-382A-1 x AAW03563

Align seg 1/1 to: AAW03563 from: 1 to: 92

4 TTGACCGAGCTGGAGAAAGCCTTGAATCTATCATACAGCTCTACACAA 53  
:::|||||  
1 MetThrLysLeuGluAspHisLeuGluGlyIleAsnIlePheHisG 17  
54 GTACTCCCTGATTAAGGGAATTTCCATGCCGCTACAGGATGACCTGA 103  
:|||||  
17 nTysSerValArgValGlyHisPheAspThrLeuAsnLysArgGluLeu 34  
104 AGAATTTGTAGAGACCGAGTGTCTCAGTATATCAGG..... 141  
|||||  
34 ysgLleuIleThrLysGluLeuProLysThrLeuLnsnThrLysasp 50  
142 AAAAAAGGTGCAGACGCTGTCTCAAGAGTTGGATATCAACACGATGG 191  
:::|  
51 GluProThrIleAspLysIlePheGlnAspLeuAspLysAsp 67  
192 TGCAGTTAACTTCAGGAGTTCCTCATTTGATGATAAAGAGCGCTGG 241  
|||||  
67 yAlaValSerPheGluGluPheValIleValSerArgValLeuLysT 84  
242 CAGCCCAAAAAAGCCATGAGAA 267  
:::|||||  
84 hAlaHisIleAspIleHisLysGlu 92

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:AAW03564

seq\_documentation\_block:

ID AAW03564 standard; Protein; 92 AA.

AAW03564;

01-MAY-1997 (first entry)

Calcium binding protein CAAFL.

Calcium binding protein; human; amniotic fluid; S100 protein family;  
intracellular signal transduction; squamous epithelial cell; neutrophil;  
macrophage; cancer; cancerous lesion; inflammation; neoplasia; cervix;  
squamous cell carcinoma; skin; oesophagus; CAAFL; lung; blood disease.

Homo sapiens.

EP731166-A2.

11-SEP-1996.

04-DEC-1995; 95EP-0119045.

06-MAR-1995; 95JP-0070468.

06-MAR-1995; 95JP-0045564.

(HITO/) HITOMI J.

(TOFU) TONEN CORP.

Hitomi J, Kimura T, Yamaguchi K, Yamamura T;

WPI; 1996-403989/41.

N-PSDB; AAT39346.

New human or bovine calcium binding protein and related nucleic acid  
- is a marker for inflammation, neoplasia, skin and blood diseases

Claim 1; Page 24; 36pp; English.

This sequence represents the CAAFL calcium-binding protein isolated from  
human amniotic fluid. CAAFL belongs to the S100 protein family, which  
includes calyculin, MRP8, and MRP14. Intracellular calcium ion  
concentration is one of the key factors for intracellular signal  
transduction. The calcium signals are transduced by various  
calcium-binding proteins, such as this protein. CAAFL is normally  
expressed in squamous epithelial cells, neutrophils and macrophages, but  
atypical epithelial cells are negative for CAAFL and overexpression is  
observed in several types of cancer cells and neutrophils/macrophages

CC infiltrating cancerous lesions. Detection of CAM1 (using antibodies in  
CC usual immunossays) can be used to diagnose (or monitor) inflammation,  
CC neoplasia (particularly squamous cell carcinoma of the skin, oesophagus,  
CC lung and cervix), and skin and blood diseases.  
XX

Sequence 92 AA:

alignment\_scores:  
Quality: 158.00 Length: 96  
Ratio: 2.508 Gaps: 2  
Percent Similarity: 65.625 Percent Identity: 38.542

alignment\_block:

US-09-806-382A-1 x AAW03564 ..

Align seg 1/1 to: AAW03564 from: 1 to: 92

```
4 TTGACCGAGCTGGAGAAAGCCTTGAACCTGATCATGACGCTACACAA 53
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1 MetThrLysLeuGluGlnHisLeuGluGlyIleValAsnIlePheHisG1 17

54 GTACTCCCTGATTAAGGGGAATTTCCATGCCGTACAGGATGACCTGA 103.
|||||::: |||||::: |||::: |||::: |||::: |||::: |||
17 nTyrSerValArgLysGlnHisPheAspThrLeuSerLysGlyIleuL 34

104 AGAATTCGTAGACCGAGTGTCTCATATATACG..... 141
||:::||||| |||::: |||::: |||::: |||::: |||::: |||
34 ysgInleuLeuThrLysGluLeuAlaAsnThrIleLysAsnIleLysAsp 50

142 AAAAGGGGTGCAGACGCTGTGTTCAAGAGTTGATATCAACACTGATG 191
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 LysAlaValIleAspGluIlePheGlnGlyLeuAspAlaAsnGlnAspG1 67

192 TGCAGTTACTTCAGAGAGTTCCTCATTCGTGATTAAGATGGCGCTG 241
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
67 uGlnValAspPheGlnGluPheIleSerLeuVal.....AlaI 80

242 CAGCCCAAAAAAGCCATGAAGAAAGCCACAAGAG 279
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 leAlaLeuLysAlaAlaHisTyrHisThrHisLysGlu 92
```

seq\_name: /SIDSI/gcgdata/hold-geneseq/genesep-emb1/AA1997.DAT:AAW24137

seq\_documentation\_block:

ID AAW24137 standard; Protein: 92 AA.

XX AAW24137:

XX 28-JAN-1998 (first entry)

XX Human chemotactic cytokine I.

XX chemotactic cytokine; tumour; autoimmune disease; antagonist;

XX agonist.

XX Homo sapiens.

XX W09723640-A1.

XX 03-JUL-1997.

XX 26-DEC-1995; 95WO-US16871.

XX 26-DEC-1995; 95WO-US16871.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Alfonso P. Genitz R. Ni J, Su JY, Yu G;

XX WPI; 1997-351075/32.

XX N-PSDB; MAT85774.

XX.

PT DNA encoding chemotactic cytokine I - used to treat, e.g. tumours,  
PT chronic infection, leukaemia, etc.  
XX  
XX  
PS Claim 12; Pages 48-49; 64pp; English.

CC This is a human chemotactic cytokine I polypeptide. The encoding  
CC polynucleotide, along with a vector and a host cell can be used for the  
CC recombinant production of the chemotactic cytokine. Cytokine agonists  
CC and antagonists can be used for the treatment of a patient requiring a  
CC chemotactic cytokine I and for the treatment of a patient requiring the  
CC inhibition of a chemotactic cytokine I polypeptide, respectively. The  
CC chemotactic cytokine is used to treat tumours, chronic infection,  
CC leukaemia and T-cell mediated autoimmune diseases.

XX Sequence 92 AA:

alignment\_scores:  
Quality: 158.00 Length: 96  
Ratio: 2.508 Gaps: 2  
Percent Similarity: 65.625 Percent Identity: 38.542

alignment\_block:

US-09-806-382A-1 x AAW24137 ..

Align seg 1/1 to: AAW24137 from: 1 to: 92

```
4 TTGACCGAGCTGGAGAAAGCCTTGAACCTGATCATGACGCTACACAA 53
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
1 MetThrLysLeuGluGlnHisLeuGluGlyIleValAsnIlePheHisG1 17

54 GTACTCCCTGATTAAGGGGAATTTCCATGCCGTACAGGATGACCTGA 103
|||||::: |||||::: |||::: |||::: |||::: |||::: |||
17 nTyrSerValArgLysGlnHisPheAspThrLeuSerLysGlyIleuL 34

104 AGAATTCGTAGACCGAGTGTCTCATATATACG..... 141
||:::||||| |||::: |||::: |||::: |||::: |||::: |||
34 ysgInleuLeuThrLysGluLeuAlaAsnThrIleLysAsnIleLysAsp 50

142 AAAAGGGGTGCAGACGCTGTGTTCAAGAGTTGATATCAACACTGATG 191
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 LysAlaValIleAspGluIlePheGlnGlyLeuAspAlaAsnGlnAspG1 67

192 TGCAGTTACTTCAGAGAGTTCCTCATTCGTGATTAAGATGGCGCTG 241
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
67 uGlnValAspPheGlnGluPheIleSerLeuVal.....AlaI 80

242 CAGCCCAAAAAAGCCATGAAGAAAGCCACAAGAG 279
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 leAlaLeuLysAlaAlaHisTyrHisThrHisLysGlu 92
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2002, 05:21:34 ; Search time 66.65 Seconds  
(without alignments)  
1039.290 Million cell updates/sec

Title: US-09-806-382A-1

Perfect score: 282  
Sequence: 1 atgttgaccgagctgagaa.....aagaagccacaagagtag 282

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents.NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCtus.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	282	100.0	408	1	US-08-385-241-2
2	282	100.0	418	1	US-07-987-272A-15
3	145.8	51.7	4195	1	US-08-200-016-1
4	142.8	50.6	433	1	US-07-987-272A-13
5	65.6	23.3	429	2	US-08-568-310D-1
6	65.6	23.3	429	4	US-09-270-455-1
7	50.8	18.0	273	3	US-08-794-000-3
8	48.4	17.2	441	4	US-08-568-310D-12
9	48.4	17.2	441	4	US-09-270-455-12
10	40.8	14.5	571	1	US-08-385-241-4
11	40	14.2	303	1	US-08-190-560-1
12	40	14.2	303	1	US-08-469-277-1
13	40	14.2	303	2	US-08-468-946-1
14	40	14.2	303	2	US-08-468-942-1
15	40	14.2	579	1	US-08-190-560-3
16	40	14.2	579	1	US-08-469-277-3
17	40	14.2	579	2	US-08-468-946-3
18	40	14.2	579	2	US-08-468-942-3
19	38.2	13.5	10952	1	US-08-602-036A-1
20	38.2	13.5	10952	2	US-08-502-374A-1
21	38.2	13.5	10952	2	US-08-642-407A-1
22	37.4	13.3	4440	1	US-08-200-016-4
23	36.8	13.0	452	1	US-07-662-198B-1
24	36.8	13.0	452	1	US-08-322-742-1
25	33.6	11.9	7218	1	US-08-232-463-14
26	31.8	11.3	2179	2	US-08-551-356-3
27	31.8	11.3	2179	5	PCT-US93-12687-3

C	28	31.8	11.3	5943	1	US-08-206-176-1	Sequence 1, Appl
	29	30	10.6	1201	1	US-09-048-889-8	Sequence 8, Appl
	30	29.8	10.6	1101	3	US-08-895-707-5	Sequence 5, Appl
	31	29.8	10.6	1158	3	US-08-895-707-8	Sequence 8, Appl
	32	29.8	10.6	1179	1	US-08-706-539-3	Sequence 3, Appl
	33	29.8	10.6	1179	4	US-09-027-007-3	Sequence 3, Appl
	34	29.8	10.6	1188	1	US-08-706-539-2	Sequence 2, Appl
	35	29.8	10.6	1188	4	US-09-027-007-2	Sequence 2, Appl
	36	29.8	10.6	1311	1	US-08-706-539-4	Sequence 4, Appl
	37	29.8	10.6	1311	4	US-09-027-007-4	Sequence 4, Appl
	38	29.8	10.6	1680	1	US-08-706-539-1	Sequence 1, Appl
	39	29.8	10.6	1680	1	US-08-522-166-1	Sequence 1, Appl
	40	29.8	10.6	1680	1	US-08-488-382A-1	Sequence 1, Appl
	41	29.8	10.6	1680	1	US-08-385-142-1	Sequence 1, Appl
	42	29.8	10.6	1680	2	US-08-480-912-1	Sequence 1, Appl
	43	29.8	10.6	1680	4	US-09-027-007-1	Sequence 1, Appl
	44	29.8	10.6	1692	1	US-08-485-859-1	Sequence 1, Appl
	45	29.8	10.6	1951	3	US-08-895-707-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-385-241-2  
; Sequence 2, Application US/08385241  
; Patent No. 5776348  
; GENERAL INFORMATION:  
; APPLICANT: Selenaut Ph.D., Jeremy D.  
; APPLICANT: Orme-Johnson Ph.D., William H.  
; APPLICANT: Dretler M.D., Stephen P.  
; APPLICANT: Asakura M.D., Hirokaka  
; TITLE OF INVENTION: SYSTEM AND METHOD FOR INHIBITING  
; TITLE OF INVENTION: FORMATION OF CRYSTALLINE STRUCTURES THAT INCLUDE STRUTITE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Choate, Hall & Stewart  
; STREET: 53 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2891  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/385,241  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Herschbach Ph.D., Brenda M.  
; REGISTRATION NUMBER: P-39,223  
; REFERENCE/DOCKET NUMBER: 492611-000 (MIT6915)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-5175  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 408 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; IMMEDIATE SOURCE:  
; CLONE: hmrp-8 cDNA  
; US-08-385-241-2  
Query Match 100.0%; Score 282; DB 1; Length 408;

Best Local Similarity 100.0%; Pred. No. 4.2e-87;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttgaccgagctggggaagccttgaactctatcatcgaactctccacaagtactcc 60  
Db 57 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATCGACGCTACCAAGTACTCTCC 116  
QY 61 ctgataaagggaattccatccgctctacaggaatgacctgaagaattgctagaagcc 120  
Db 117 CTGATTAAGGGGAAATTTCCATGCCGCTCTACAGGATGACTGGAAGAAATTCCTAGAGACC 176  
QY 121 gagtgtccctcaatatacagaagaaagggtgcagacgtctggttcaagaagtggatatac 180  
Db 177 GAGTGTCTCTCAGTATATCAGGAAAAAGGTGCAGACGCTGTGTTCAAGAAGTTGATATC 236  
QY 181 aacctatggtgcaagttaacttcccaagagttctcattctgtgtataaagaatggcgctg 240  
Db 237 AACACTGATGTGTCAGTAACTTCCAGGAGTTCCTCATTTCTGTGATTAAGATGGCGGTG 296  
QY 241 gcaagcccaaaaaaagccatgaagaagccacaagaagtag 282  
Db 297 GCAGCCCAAAAAAGCCATGAAGAAAGCCACAAAGATAG 338

RESULT 2  
US-07-987-272A-15  
; Sequence 15, Application US/07987272A  
; Patent No. 5731166  
; GENERAL INFORMATION:  
; APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M  
; TITLE OF INVENTION: No. 5731166el Chemotactic Factor  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Cushman Darby & Cushman  
; STREET: 1100 New York Avenue, N.W., Ninth Floor, East Tower  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/987,272A  
; FILING DATE: 05-Mar-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PK 2127  
; FILING DATE: 05-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PK 4463  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brinkman, David W  
; REGISTRATION NUMBER: 20,817  
; REFERENCE/DOCKET NUMBER: DMB/1925/200259  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861 3000  
; TELEFAX: 202-822 0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 418 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 57..218  
; FEATURE:

; NAME/KEY: CDS  
; LOCATION: 57..335  
US-07-987-272A-15

Query Match 100.0%; Score 282; DB 1; Length 418;  
Best Local Similarity 100.0%; Pred. No. 4.3e-87;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttgaccgagctggggaagccttgaactctatcatcgaactctccacaagtactcc 60  
Db 57 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATCGACGCTACCAAGTACTCTCC 116  
QY 61 ctgataaagggaattccatccgctctacaggaatgacctgaagaattgctagaagcc 120  
Db 117 CTGATTAAGGGGAAATTTCCATGCCGCTCTACAGGATGACTGGAAGAAATTCCTAGAGACC 176  
QY 121 gagtgtccctcaatatacagaagaaagggtgcagacgtctggttcaagaagtggatatac 180  
Db 177 GAGTGTCTCTCAGTATATCAGGAAAAAGGTGCAGACGCTGTGTTCAAGAAGTTGATATC 236  
QY 181 aacctatggtgcaagttaacttcccaagagttctcattctgtgtataaagaatggcgctg 240  
Db 237 AACACTGATGTGTCAGTAACTTCCAGGAGTTCCTCATTTCTGTGATTAAGATGGCGGTG 296  
QY 241 gcaagcccaaaaaaagccatgaagaagccacaagaagtag 282  
Db 297 GCAGCCCAAAAAAGCCATGAAGAAAGCCACAAAGATAG 338

RESULT 3  
US-08-200-016-1  
; Sequence 1, Application US/08200016  
; Patent No. 5614397  
; GENERAL INFORMATION:  
; APPLICANT: Weissman, Irving  
; TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR MODULATING  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road, Suite 100  
; CITY: Menlo Park  
; STATE: California  
; COUNTRY: US  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/200,016  
; FILING DATE: 22-FEB-1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rae-Venter, Barbara  
; REGISTRATION NUMBER: 32,750  
; REFERENCE/DOCKET NUMBER: 06037/003001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 854-5277  
; TELEFAX: (415) 854-0875  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4195 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-200-016-1

```

QY 126 tctcagatatactcgaaaaaggtgcagacgtctggtcctcaaaagattggatatcaaac 185
Db 2316 TCTCGAATCTTGACAGAAAAAGGCGTGCAGAGCTGTGTTCAAAAGATTGGATATCAACAC 2375
QY 186 tgaatgtgcagtttaacttcacagagattccatctctgtgtatataagatggtgcgcgc 245
Db 2376 TGATGTGTGTCAGTTAACTTCCAGAGAGTTCCTCATCTGTGTGATAAAGATGCGCGTGCAGC 2435
QY 246 ccacaaaaaacgcattgaagaagccacaaagagtag 282
Db 2436 CCACAAAAAAGCCATTGAAGAAAGCCACAAAGTAG 2472

RESULT 4
US-07-987-272A-13
: Sequence 13 Application US/07987272A
: Patent No. 5731166
: GENERAL INFORMATION:
: APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M
: TITLE OF INVENTION: No. 5731166el Chemotactic Factor
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Cushman Darby & Cushman
: STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
: City: Washington
: STATE: D. C.
: COUNTRY: USA
: ZIP: 20005-3918
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/987, 272A
: FILING DATE: 05-MAR-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PK 2127
: FILING DATE: 05-FEB-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PK 4463
: FILING DATE: 05-SEP-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Brinkman, David W
: REGISTRATION NUMBER: 20,817
: REFERENCE/DOCKET NUMBER: DMB/1925/200259
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-861 3000
: TELEFAX: 202-822 0944
: TEXT: 6714627 CUSH
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 433 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 52..318
: US-07-987-272A-13

Query Match 51.7%; Score 145.8; DB 1; Length 4195;
Best Local Similarity 95.5%; Pred. No. 5.6e-40;
Matches 150; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1 atgttaccggcgtgagaaagccttgacatcatcatcagtcacgtctaccacaagtactcc 60
Query Match 50.6%; Score 142.8; DB 1; Length 433;
Best Local Similarity 71.1%; Pred. No. 1.8e-39;
Matches 189; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```

```

Db      52  ATGCCGTTGAACTBGGAGAAGSCCTTAGCAACTCATGTATGTCTACACAAATTATTCC 1111
QY      61  ctgataaagggaatttccagtcgcgctctacaggagtacctaagaatgtctgaagac 120
Db      112  AATATACAAGGAATACACATGCGCCTCTACAAGAATGACTCTCAAGAAATAATGTCACACT 171
QY      121  gagggtcctcagtatatacaggaataagggtgcagagtcgtggttcaaggatgtgatac 180
Db      172  GAGTCTCTCTCAGTTTGTCAGAAATATAAATATCGAAAACTTGTTCAAGAAATGAGACT 231
QY      181  aacctgtatgtgcagtttaacttcacaggagtctcctcaattctgtgtataaagatggcg 240
Db      232  AATAATGCAATGCAATTAACTTTGAGGAGAGTTCCTTGCGATGTGATATAAAGTGGGTG 291
QY      241  gaagcccaaaaaaaggccatgaaga 266
Db      292  GCATCTCACAAAGACAGCCACAGA 317

RESULT 5
US-08-568-310D-1
; Sequence 1, Application US/08568310D
; Patent No. 5976832
GENERAL INFORMATION:
APPLICANT: HITOMI, JIRO
APPLICANT: YAMAGUCHI, KEN
APPLICANT: YAMAMURA, TOKUJIRO
APPLICANT: KIMURA, TATSUJI
TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESS: WYATT, GERBER, MELLER & O'ROURKE
STREET: 99 PARK AVENUE
STREET: 6th FLOOR
CITY: NEW YORK CITY
STATE: NEW YORK
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB
MEDIUM TYPE: STORAGE
COMPUTER: IBM-PC COMPATIBLE
OPERATING SYSTEM: PC-DOS 6.2
SOFTWARE: WORDPERFECT 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568, 310D
FILING DATE: DECEMBER 6, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 7-70468 and 7-45564 (both Japan)
FILING DATE: 3/6/95 and 3/6/95, respectively
ATTORNEY/AGENT INFORMATION:
NAME: KLEIN, MILTON
REGISTRATION NUMBER: 27101
REFERENCE/DOCKET NUMBER: 3316
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)953-3350
TELEFAX: (212)953-3352
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 429
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 1:
RELEVANT RESIDUES IN SEQ ID NO: FROM 1 TO 429
US-08-568-310D-1

```

```

? LENGTH: 429
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? PUBLICATION INFORMATION:
? RELEVANT RESIDUES IN SEQ ID NO: 1 : FROM 1 TO 429
? US-09-270-455-1

```

	Query Match	Score 50.8%; DB 3; Length 273;
	Best Local Similarity	52.6%; Pred. No. 4e-08;
Matches 143; Conservative 0; Mismatches 117; Indels 12; Gaps 1.		
QY	7 accaggtctggagaagaaccttgtaactctatcatctgtagcttaccacaagtctcccgata	66
Db	1 ACCAAGCTGGAGGACCACTCGTAGGGCCATCATCAACTTCCACCAGTAICTGTGTCGG	60
QY	67 aagggaatttcacatgcctctcaagaagatgaccttgaaagaatttgtctagaagcagagtgt	126
b	61 CTGGGGCACTATTGAACNCCCTGCATCAGAAGGGGAGCTGTGAAGCAGTGATCACCAAGGAGCTG	120



```

QY      127 cctaagatatacggaaaaagggtgcagaacgcttg-----ttcaaaagttg 174
           || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 CCCAACACCTGTGAAGAAACCAAGAACCAGGSCACCATTTGCAMAGATCTTCCAGAACCTG 180
QY      175 gatatcacactgatgtgtaagtltaacttcacagagattcccatctcgtagtataaatg 234
           || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181 GATGCCAACACGAGATGAGCAGGAGTGTCCTTCAAGSGAGTTGTGGTGGTGGTGGACGATGCTG 240
QY      235 ggctgtgagagcccacaataaaaagccatgaaga 266
           | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      241 CTGATCACAGCCCATGACAACATCCACAAGGA 272

RESULT      8
US-08-568-310D-12
; Sequence 12, Application US/08568310D
; Patent No. 5976832
; GENERAL INFORMATION:
; APPLICANT: HITOMI, JIRO
; APPLICANT: YAMAGUCHI, KEN
; APPLICANT: YAMAMURA, TOKUJIRO
; APPLICANT: KIMURA, TATSUJI
; TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE
; STREET: 99 PARK AVENUE
; CITY: NEW YORK CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB
; MEDIUM TYPE: STORAGE
; COMPUTER: IBM-PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS 6.2
; SOFTWARE: WORDPERECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,310D
; FILING DATE: DECEMBER 6, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7-70468 and 7-45564(both Japan)
; FILING DATE: 3/6/95 and 3/6/95,respectively
; ATTORNEY/AGENT INFORMATION:
; NAME: KEVIN, MILTON
; REGISTRATION NUMBER: 27101
; REFERENCE/DOCKET NUMBER: 3316
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)953-3350
; TELEFAX: (212)953-3352
; INFORMATION FOR SEQ ID NO: 12:
; LENGTH: 441
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 12:
; RELEVANT RESIDUES IN SEQ ID NO: FROM 1 TO 441
US-08-568-310D-12

Query Match      17.2%; Score 48.4; DB 2; Length 441;
Best Local Similarity 48.2%; Pred.No. 3.5e-07;
Matches 136; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY      1 atgttacccgagtgtgagaagaccttgaaacttatcatcagcagctctaccacaagtactcc 60
           | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      19 AAGTGTCAAAATTTTAAGAGCATCTGGAGGGAATGTCAATATCTTCCACCAATACTCA 78

```

Oy	61	cgtataaaggaggaatttcattgccgtgtacaaggatgtaactgaaagaattgctagagacc	120
Db	79	gttcggaaaggsggcatttttgacaccctctttaaagggtagctggaagcagcgtttacaaag	138
Oy	121	gagtgccttcgatatcatcagaanaaagggtgcagagcttggtgtcaagaagttggaatc	180
Db	139	gagcttgcaaacaccatcacaagaatttcacaagatataaactgtcattgtatgaaattttcca	198
Oy	181	aacactgtatgtcagtttaacttccaagagttctcatctctggtataaagaatgggcgtg	240
Db	199	ggccctggatgcttaattcaagaatgaacagctgcacttttcagaaattcatatccctgtgacc	258
Oy	241	gcagcccaaaaaaaggcattgaagaagaagccacaagaagttag	282
Db	259	atttgccctgaagagctgcccatattccacaccaccaaagaatgag	300

```

RESULT      9
US-09-270-455-12
; Sequence 12, Application US/09270455
Patent No. 631367
GENERAL INFORMATION:
APPLICANT: HITOMI, JIRO
APPLICANT: YAMAGUCHI, KEN
APPLICANT: YAMAMURA, TOKUJIRO
APPLICANT: KIMURA, TATSUJI
TITLE OF INVENTION: NOVEL CALCIUM-BINDING PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: WYATT, GERBER, MELLER & O'ROURKE
STREET: 99 PARK AVENUE
STREET: 6th FLOOR
CITY: NEW YORK CITY
STATE: NEW YORK
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH, 720 KB
MEDIUM TYPE: STORAGE
COMPUTER: IBM-PC COMPATIBLE
OPERATING SYSTEM: PC-DOS 6.2
SOFTWARE: WORDPERFECT 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/270.455
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/568,310
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KLEIN, MILTON
REGISTRATION NUMBER: 27101
REFERENCE/DOCKET NUMBER: 3316
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)953-3350
TELEFAX: (212)953-3352
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 441
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 12: FROM 1 TO 441
US-09-270-455-12

```

Qy	5	tgacgcgagcttgagagaagaagcccttgaaactctatcatcgagcgtctaccagaaghuacccctga	64
Db	57	TGTGCACCTTGGTGAAGCGCAACATAGAGACCATCATCAACACCTTCCACCAATACTCTGTGA	116
Qy	65	taaaagggaattccatcgtccgtctacagggatgaacctgaagaanaattgctagagaccgagt	124
Db	117	AGCTGGGGCCACCCAGACACCTTGTAACCGAGGGGGGAATTCAAGAGCTGGTGGCAAAAATATC	176
Qy	125	gtccttcagtatatacaagg-----aaaaggtgcgcgaagctctggttcaag	165
Db	177	TGCAAAATTTTCTCAAGAGAGGAGATAAAGATGAAGAAAGGTGCATAGAAACATCTATGGAGG	236
Qy	170	agttgagatacaacacatbatgtgtgcagtttaacttccagagattccatctcgtgtga	229
Db	237	AACCTGACACAAATATGCAGACAAAGACGCTTGAGCTTCGAGAGATTCATCATGCTGATGGCGA	296
Qy	230	agatgggcgttgccagcccaaaaaaagccatgaag	265
Db	297	GGCTAACCTGGGCTCCACACAGAAAGATGACGCGAGG	332

Query Match <sub>h</sub>	14.5%;	Score 40.8;	DB 1;	Length 571;
Best Local Similarity	50.4%;	Pred. No. 0.00016;		
Matches 139; Conservative	0;	Mismatches 122;	Indels 15;	Gaps 1

Qy	5	tgacgcgagcttgagagaagcccttgacctatcatcgcgagcttaccacaaghuacctctga	64
Db	57	TGTGCACCTTGGTAAGCGAACATATAGAGACCATCATCAACACCTTCCACCAATACTCTGTGA	116
Qy	65	taaaagggaattccatcgtccgtctacagggatgaacctgaagaanaattgctagagaccgagt	124
Db	117	AGCTGGGGCCACCCAGACACCTCGTAACCGAGGGGGAAATTCAAAGACCTGGTGGCAAAAATATC	176
Qy	125	gtccttcagtatatacaagg-----aaaaggtgcgcgaagctctggttcaag	165
Db	177	TGCAAAATTTCTCAAGAGGAGGAGATAAGATGAAGAAAGGTGCATATGAACACATCTGAGAG	236
Qy	170	agttgatatcaacacactbatgtgtgcagtttaacttccagaggttccatctctgtgtga	229
Db	237	AACCTGACACAAATATGCAGACAAAGACGCTTGAGCTTCGAGAGATTCATCATGCTGATGGCGA	296
Qy	230	agatgggcgtggcagcccaaaaaaagccatgaag	265
Db	297	GGCTAACTGGGCTCCACACAGAAAGATGACGCGAGG	332

Query Match	14.2%;	Score 40;	DB 1;	Length 303;
Best Local Similarity	49.1%;	Pred. No. 0.00021;		
Matches 106;	Conservative	0;	Mismatches 110;	Indels 0;
				Gaps 0.

QY 1 atgtgaccgagctgtgagaagccttgaaacctatcatcagcgtctaccaaagtactcc 60  
||| | | ||||| ||| | | | | ||||| |||||  
Db 1 ATGGCTGCCCCCTCTGGAGAGAGGCCCTGGATGTGATGGTGTCCACCTTCCACAAGTACTCG 60

```

QY 121 gattgctccatgatatcaagaaagaggtgcagacgtctggttcaagaatggtatc 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GACTGTGCCACGCTTCTTGGGGAAAAAGACAGATGAAGTCGTCTTCCAGAACGTGATGAC 180

QY 181 aacactgatgtgtgcagtaacttcacagaggtcttc 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 AACCTTGACACGCAACAGGGAGCAACGAGGTGGACCTTC 216

RESULT 13
US-08-468-946-1
Sequence 1, Application US/08468946

```

```

1      RESULT 13
2      US-08-468-946-1
3      : Sequence 1/, Application US/08468946
4      : Patent No. 5843686
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: Zain, Sayeeda
9      : APPLICANT: Lukhanidin, Eugene
10     : TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
11     : TITLE OF INVENTION: THE MTS-1 GENE
12     : NUMBER OF SEQUENCES: 9
13     : CORRESPONDENCE ADDRESS:
14     : ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
15     : STREET: 400 Garden City Plaza
16     : CITY: Garden City
17     : STATE: New York
18     : COUNTRY: United States
19     : ZIP: 11530
20     :
21     : COMPUTER READABLE FORM:
22     : MEDIUM TYPE: Floppy disk
23     : COMPUTER: IBM PC compatible
24     : OPERATING SYSTEM: PC-DOS/MS-DOS
25     : SOFTWARE: PatentIn Release #1.0, Version #1.25
26     : CURRENT APPLICATION DATA:
27     : APPLICATION NUMBER: US/08/468, 946
28     : FILING DATE: 06-JUN-1995
29     : CLASSIFICATION: 435
30     :
31     : PRIOR APPLICATION DATA:
32     : APPLICATION NUMBER: US 08/190,560
33     : FILING DATE: 31-JAN-1994
34     : ATTORNEY/AGENT INFORMATION:
35     : NAME: Digiglio, Frank S.
36     : REGISTRATION NUMBER: 31,346
37     : REFERENCE/DOCKET NUMBER: 78792Y
38     : TELECOMMUNICATION INFORMATION:
39     : TELEPHONE: (516) 742-4343
40     : TELEFAX: (516) 742-4366
41     : TELEX: 230 901 SANS UR
42     : INFORMATION FOR SEQ ID NO: 1:
43     : SEQUENCE CHARACTERISTICS:
44     : LENGTH: 303 base pairs
45     : TYPE: nucleic acid
46     : STRANDEDNESS: double
47     : TOPOLOGY: linear
48     :
49     : MOLECULE TYPE: CDNA
50     :
51     : US-08-468-946-1

```

Query Match	14.2%	Score 40:	DB 2:	length 303;
Best Local Similarity	49.1%	Prod. No.	0.00021:	
Matches 106;	Conservative 0;	Mismatches 110;	Indels 0;	Gaps
Qy	1	algtttacccagagctltgagaaagccctltgaactatataatcagcttctacccaagtactcc	60	
Db	1	ATGGCGTCGCCCTCTGGAGAGAGCCCTGGATGTATGTGTGCCACCTTCACAAAGTACTCG	60	
Qy	61	ctataaagggaatttccatcagccgtctacaaggatgaacttgaagaattgtctagagacc	120	
Db	61	GGCAAGAGAGGGTGACAAAGTTCAAGCTCAACCAAGTCAGAGTAAAGGAGCGTGTACCCGG	120	
Qy	121	gagttgctccagtatatacagaaaagagtgagcagcgtctggttcaagaagtttgatatc	180	
Db	121	GAGCTGCCACGCTTCTTGGGGAGAAAGACAGATGSAAGCTCTTTCCAGAGAGCTGATGAGC	180	
Qy	181	aacactgatgtgtgacagtttaacttccaggaagttcc	216	



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2002, 05:22:44 ; Search time 298.43 Seconds

(without alignments)  
1622.390 Million cell updates/sec

Title: US-09-806-382A-1

Perfect score: 282

Sequence: 1 atgttgaccgagctgtgagaa.....aagaagaaccacaagagtag 282

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: N\_Geneseq\_032802.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	282	100.0	282	21	AAC81808	Human S100A8 CDNA.
2	282	100.0	282	21	AAA11969	Human calcium-bind
3	282	100.0	408	19	AAV34697	Human calprotectin
4	282	100.0	944	22	AAK51509	Human polynucleoti
5	282	100.0	1144	22	AAH73234	Human cervical can
6	282	100.0	1144	22	AAH73234	Human cervical can
7	278.4	98.7	2329	22	AAH72611	Human cervical can
8	249.6	88.5	645	22	AAK52493	Human polynucleoti
9	230.4	81.7	573	22	AAH68950	Human cervical can

10	218.4	77.4	497	22	AAH70917	Human cervical can
11	218.4	77.4	598	22	AAH71815	Human cervical can
12	218.4	77.4	625	22	AAH70613	Human cervical can
13	199	70.6	381	22	AAH71977	Human cervical can
14	199	70.6	512	22	AAH69112	Human cervical can
15	191	67.7	287	22	AAH71248	Human cervical can
16	190	67.4	333	22	AAH69823	Human cervical can
17	145.8	51.7	4195	18	AAH68321	Human multitrug re
18	145.8	51.7	4195	22	AAH68321	Human multitrug re
19	144.8	51.3	254	16	AAH70087	Nucleotide sequenc
20	142.8	50.6	433	13	AAQ23343	Human gene signatu
21	65.6	23.3	429	17	AAH39345	Murine CP-10. Mus
22	61.4	21.8	395	21	AAH29815	Calcium binding pr
23	57	20.2	305	22	AAH71963	Bovine EN-RAGE cDN
24	57	20.2	307	22	AAH69098	Human cervical can
25	57	20.2	358	22	AAH70761	Human cervical can
26	55.8	19.8	273	18	AAH62569	Human S100A5 CDNA.
27	50.8	18.0	273	18	AAH62569	DNA encoding compo
28	50.6	17.9	93	22	AAH71206	Human cervical can
29	50.6	17.8	141	22	AAH72274	Human cervical can
30	50.2	17.8	342	22	AAH54709	Nucleotide sequenc
31	48.4	17.2	440	17	AAH39346	Calcium binding pr
32	48.4	17.2	479	18	AAH85774	Human chemotactic
33	47.6	16.9	279	21	AAH81812	Human S100A12 CDNA
34	46.2	16.4	234	19	AAV47616	Nucleotide sequenc
35	43.2	15.3	573	22	AAH62023	Human foetal liver
36	43.2	15.3	573	22	AAH62023	Probe #7960 for ge
37	43.2	15.3	573	22	AAH10340	Human brain expres
38	43.2	15.3	573	22	AAH36240	Human bone marrow
39	43.2	15.3	573	22	AAH41960	Probe #10646 used
40	43	15.2	524	22	AAH33227	Human colon cancer
41	42.4	15.0	407	23	AAH57328	cDNA #4 encoding p
42	42.2	15.0	712	22	AAH99198	Human protein enco
43	41.8	14.8	273	21	AAH81806	Human S100A6 CDNA.
44	41.8	14.8	452	22	AAH54717	Nucleotide sequenc
45	41.8	14.8	704	20	AAH24438	Human bladder tumo

## ALIGNMENTS

RESULT 1	
ID AAC81808	standard; cDNA; 282 BP.
XX	
AC AAC81808;	
XX	
DT 22-FEB-2001	(first entry)
XX	
DE Human S100A8 CDNA.	
XX	
KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;	
KW calcium-binding protein; calcium homeostasis; cardiac muscle;	
KW pumping capacity; myocardial cell; systolic calcium ion release;	
KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;	
KW valve defect; ss.	
XX	
OS Homo sapiens.	
XX	
PN DE19915485-A1.	
XX	
PD 19-OCT-2000.	
XX	
PF 07-APR-1999;	99DE-1015485.
XX	
PR 07-APR-1999;	99DE-1015485.
XX	
PA (KATU/) KATU5 H A.	
XX	
PI (REMP/) REMPPIS A.	
XX	
PI Katus HA, Remppis A;	
XX	
DR WPI; 2000-673510/66.	

DR P-PSDB: AAB45538.  
XX  
XX Composition containing S100 protein, corresponding nucleic acid or  
PT vector, useful for treating cardiomyopathy and cardiac insufficiency -  
XX  
XX Claim 36; Page 15; 36pp; German.  
XX  
CC This invention describes a novel composition for treating primary or  
CC secondary cardiomyopathy or cardiac insufficiency contains at least one  
CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or  
CC fragments, or a gene transfer vector containing (II), optionally  
CC formulated with auxiliaries and/or carriers. (I) are calcium-binding  
CC proteins involved in calcium homeostasis, so their overexpression in  
CC cardiac muscle will improve pumping capacity (and overall capacity) of  
CC the heart. In cultured myocardial cells they increase the contraction and  
CC relaxation rates associated with increased systolic calcium ion release  
CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are  
CC used to treat cardiomyopathy (CMP) where inherited or caused by  
CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,  
CC diastolic CMP caused by toxic/infectious disease, cardiac disease caused  
CC by pulmonary and/or arterial hypertension, and structural disease caused  
CC by rhythm disorders or valve defects, generally any condition associated  
CC with reduced contractile force. Unlike calmodulin, which is expressed  
CC ubiquitously, (I) show tissue-specific expression and treat the  
CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac  
CC disease.  
XX  
SQ Sequence 282 BP; 90 A; 59 C; 72 G; 61 T; 0 other;  
  
Query Match 100.0%; Score 282; DB 21; Length 282;  
Best Local Similarity 100.0%; Pred. No. 1e-80;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 atgttaccgagctgagagaagccttgaactctatcatcgagcgtctccacaagtactcc 60  
DB 1 atgttaccgagctgagagaagccttgaactctatcatcgagcgtctccacaagtactcc 60  
QY 61 ctgataaagggaatttcacatcgctctacagagatgacccgaagaattgctagagacc 120  
DB 61 ctgataaagggaatttcacatcgctctacagagatgacccgaagaattgctagagacc 120  
QY 121 gagtgcctcagatatacaggaanaagggtgcagacgtctggttcaagaattgagatacc 180  
DB 121 gagtgcctcagatatacaggaanaagggtgcagacgtctggttcaagaattgagatacc 180  
QY 181 aacactgtgtgtcagtttaacttccagaggttctcattctggtgataaagatggcgctg 240  
DB 181 aacactgtgtgtcagtttaacttccagaggttctcattctggtgataaagatggcgctg 240  
QY 241 gcagcccaaaaaaagccatgagaagaagccacaaagagtag 282  
DB 241 gcagcccaaaaaaagccatgagaagaagccacaaagagtag 282  
  
RESULT 2  
AA11969  
ID AA11969 standard; DNA; 282 BP.  
XX  
XX AAA11969;  
AC  
XX  
XX 04-AUG-2000 (first entry)  
DT  
XX  
XX  
DE Human calcium-binding protein encoding DNA #1.  
XX  
XX Calcium-binding protein; granule release; calgranulin; human;  
KW vascular membrane growth; adult respiratory distress syndrome;  
KW acute myocardial infarction; ischemic reperfusion disorder;  
KW glomerulonephritis; rheumatoid arthritis; chronic bronchitis;  
KW cerebral vascular disorder; asthma; peripheral circulation disturbance;  
KW angina pectoris; hypertension; multiple sclerosis; ds.  
XX  
XX Homo sapiens.  
OS

XX  
XX WO200018970-A1.  
XX  
XX  
XX 06-APR-2000.  
XX  
XX  
XX 28-SEP-1999; 99WO-JP05302.  
XX  
XX 29-SEP-1998; 98JP-0274574.  
XX  
XX  
XX (ASAH ) ASAH KASEI KOGYO KK.  
XX  
XX Seto M, Fukuda K;  
XX  
XX WPI: 2000-293189/25.  
XX P-PSDB: AAR87636.  
DR  
XX  
XX Controlling the release of granules from cell system using activated  
PT calgranulin for screening substances for granule activating or  
PT inhibiting activity -  
XX  
XX Claim 3(1); Page 37-38; 42pp; Japanese.  
XX  
XX This invention describes a novel method for controlling the release of  
CC granules by treating a cell system to increase or decrease activated  
CC calgranulin to enhance or depress the release of granules. Calgranulin  
CC is a calcium binding protein and can be used for controlling the release  
CC of granules from a cell system e.g. those involved with the inhibition of  
CC vascular membrane growth. Vascular membrane growth is associated with  
CC e.g. adult respiratory distress syndrome, acute myocardial infarction due  
CC to ischemic reperfusion disorders, glomerulonephritis, rheumatoid  
CC arthritis, chronic bronchitis, cerebral vascular disorders, asthma,  
CC peripheral circulation disturbance, angina pectoris, hypertension and  
CC multiple sclerosis. The new method is used for screening substances for  
CC their ability to activate or inhibit the release of granules. This  
CC sequence encodes a human calcium-binding protein which is described in  
CC the method of the invention.  
XX  
SQ Sequence 282 BP; 90 A; 59 C; 72 G; 61 T; 0 other;  
  
Query Match 100.0%; Score 282; DB 21; Length 282;  
Best Local Similarity 100.0%; Pred. No. 1e-80;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 atgttaccgagctgagagaagccttgaactctatcatcgagcgtctccacaagtactcc 60  
DB 1 atgttaccgagctgagagaagccttgaactctatcatcgagcgtctccacaagtactcc 60  
QY 61 ctgataaagggaatttcacatcgctctacagagatgacccgaagaattgctagagacc 120  
DB 61 ctgataaagggaatttcacatcgctctacagagatgacccgaagaattgctagagacc 120  
QY 121 gagtgcctcagatatacaggaanaagggtgcagacgtctggttcaagaattgagatacc 180  
DB 121 gagtgcctcagatatacaggaanaagggtgcagacgtctggttcaagaattgagatacc 180  
QY 181 aacactgtgtgtcagtttaacttccagaggttctcattctggtgataaagatggcgctg 240  
DB 181 aacactgtgtgtcagtttaacttccagaggttctcattctggtgataaagatggcgctg 240  
QY 241 gcagcccaaaaaaagccatgagaagaagccacaaagagtag 282  
DB 241 gcagcccaaaaaaagccatgagaagaagccacaaagagtag 282  
  
RESULT 3  
AAV34697  
ID AAV34697 standard; cDNA; 408 BP.  
XX  
XX AAV34697;  
AC  
XX  
XX 03-SEP-1998 (first entry)  
DT  
XX

```

DE Human calprotectin subunit MRP-8 protein encoding cDNA.
XX
XX Human; MRP-8; MRP-14; calprotectin; mineral precipitate; struvite;
KM calcium phosphate; kidney stone; renal calculi; struvite stone;
XX urinary tract infection; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 57..338
FT /*tag= a
FT /product= "human MRP-8"
XX
XX US5776348-A.
XX
XX 07-JUL-1998.
XX
XX 07-FEB-1995; 95US-0385241.
XX
XX 07-FEB-1995; 95US-0385241.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Asakura H, Dretler SP, Orme-Johnson WH, Selenkut JD;
XX
XX WPI: 1998-397914/34.
XX
XX P-PSDB: AAM60177.
XX
XX Inhibiting kidney stone formation - uses the protein calprotectin
XX
XX Example 1; Columns 21-22; 19pp; English.
XX
XX This cDNA encodes a human calprotectin subunit MRP-8. This is used in
XX a method for inhibiting the formation of a mineral precipitate in a
XX solution which comprises providing a solution comprising component ions
XX of the mineral precipitate or its precursors and contacting the solution
XX with an effective amount of isolated calprotectin, or a derivative of it.
XX The method is useful for the inhibition of kidney stone formation (renal
XX calculi). Kidney stones are concentrations of inorganic and organic salts
XX that develop though crystal nucleation, aggregation and growth in the
XX kidneys, which then can block the urether and if not passed to the
XX bladder, grow and become symptomatic. Struvite stones (MgNH4PO4) are
XX commonly found after urinary tract infection. The mechanisms of this and
XX other stone formation is unclear, but some urine compositions are known
XX to inhibit formation. One of the major components of these is a protein
XX factor calprotectin, which inhibits mineral precipitation. Calprotectin
XX can also be used to raise antibodies, which in turn can be used to detect
XX the protein in samples. The levels of calprotectin found in samples can
XX be compared to levels found in normal humans, and thus assuming increased
XX calprotectin correlates to indication of kidney stone formation, this
XX procedure can be used as a diagnostic tool.
XX
XX Sequence 408 BP; 118 A; 91 C; 106 G; 93 T; 0 other;
XX
XX
XX Query Match 100.0%; Score 282; DB 19; Length 408;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-80;
XX Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 attttaccgagctgtagaagccttgactatcatcagctaccacccaagtacc 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||
XX 57 atgttgacgagctgtagaagccttgactatcatcagctaccacccaagtacc 116
XX
XX 61 ctgataaagggaatttcacatgcctctacaggaatgacctgagaatgtctagagacc 120
XX |||||||||||||||||||||||||||||||||||||||||||||||||||
XX 117 ctgataaagggaatttcacatgcctctacaggaatgacctgagaatgtctagagacc 176
XX
XX 121 gagtgtctctacgtatatacagaagaaagggtgcagcagctctggttcaagaagtgtgatac 180
XX |||||||||||||||||||||||||||||||||||||||||||||||||||
XX 177 gagtgtctctacgtatatacagaagaaagggtgcagcagctctggttcaagaagtgtgatac 236
XX
XX 181 aacactatgtgtagttaaacttcacaggaattcctcatctgtgtataaagatgggcgtg 240
XX |||||||||||||||||||||||||||||||||||||||||||||||||||

```

```

DB 237 aacactatgtgtagttaaacttcacaggaattcctcatctgtgtataaagatgggcgtg 296
XX
XX 241 gcagcccaaaaaaagccatgaagaagcccaagaagtag 282
XX |||||||||||||||||||||||||||||||||||||||||||||||||||
XX 297 gcagcccaaaaaaagccatgaagaagcccaagaagtag 338
XX
XX
XX RESULT 4
XX AAK51509
XX ID AAK51509 standard; cDNA; 944 BP.
XX
XX AAK51509;
XX
XX 06-NOV-2001 (first entry)
XX
XX DE Human polynucleotide SEQ ID NO 54.
XX
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorder; arthritis; inflammation; ss.
XX
XX OS Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX
XX 27-APR-2000; 2000US-0560875.
XX
XX 20-JUN-2000; 2000US-0598075.
XX
XX 19-JUL-2000; 2000US-0620325.
XX
XX 01-SEP-2000; 2000US-0654936.
XX
XX 15-SEP-2000; 2000US-0663561.
XX
XX 20-OCT-2000; 2000US-0693325.
XX
XX 30-NOV-2000; 2000US-0728422.
XX
XX (HYSEQ-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX Xue Au, Yang Y, Weinman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
XX
XX P-PSDB: AAM78376.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX
XX Claim 1; Page 637-638; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine/ cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
XX Sequence 944 BP; 244 A; 231 C; 223 G; 246 T; 0 other;
XX
XX
XX Query Match 100.0%; Score 282; DB 22; Length 944;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-80;

```

Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 atgttaccgagctggagaaagccttgaacctatcatcagcgtctaccagaactacc 60
Db 359 atgttaccgagctggagaaagccttgaacctatcatcagcgtctaccagaactacc 60
QY 61 ctgataaagggaatttcacgcgtctacagagatgacccgaagaattgctgagacc 120
Db 419 ctgataaagggaatttcacgcgtctacagagatgacccgaagaattgctgagacc 478
QY 121 gaggtccctcgtatatacgaagaaagggtgcagacgtctggtcacaaggttgatc 180
Db 479 gaggtccctcgtatatacgaagaaagggtgcagacgtctggtcacaaggttgatc 538
QY 181 aacactgagtgcagtttaactccagagagttcctcattcgtgtataaagatggcgctg 240
Db 539 aacactgagtgcagtttaactccagagagttcctcattcgtgtataaagatggcgctg 598
QY 241 gcagcccaaaaaaagccatgaagaagccacaagaagtag 282
Db 599 gcagcccaaaaaaagccatgaagaagccacaagaagtag 640
```

## RESULT 5

AAH73234  
ID AAH73234 standard; cDNA; 1144 BP.

AC AAH73234;

DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 4508.

DE Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

OS Homo sapiens.

PN WO200142467-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-US33312.

PR 08-DEC-1999; 99US-0169681.

PR 21-DEC-1999; 99US-0171350.

PR 14-MAR-2000; 2000US-0189315.

PR 12-MAY-2000; 2000US-0203791.

PR 09-JUN-2000; 2000US-0210600.

PR 21-JUL-2000; 2000US-0220114.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI: 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer  
and for assessing and detecting compounds for treating the cancer -

PS Claim 1; Page 986; 1051pp; English.

The invention relates to novel genes (AAH68727-AAH73383) associated with cervical cancer with cytostatic activity. The nucleic acids and encoded polypeptides are useful: to assess if a patient is afflicted with cervical cancer or has a pre-malignant condition; to monitor the progression of cervical cancer or a premalignant condition in a patient; and to select and/or assess the efficacy of a compound or therapy for inhibiting cervical cancer in a patient. The nucleic acids may also be useful for gene therapy.

Sequence 1144 BP; 347 A; 257 C; 245 G; 295 T; 0 other;

Query Match 100.0%; Score 282; DB 22; Length 1144;  
Best Local Similarity 100.0%; Pred. No. 1.9e-80;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 atgttaccgagctggagaaagccttgaacctatcatcagcgtctaccagaactacc 60
Db 515 atgttaccgagctggagaaagccttgaacctatcatcagcgtctaccagaactacc 574
QY 61 ctgataaagggaatttcacgcgtctacagagatgacccgaagaattgctgagacc 120
Db 575 ctgataaagggaatttcacgcgtctacagagatgacccgaagaattgctgagacc 634
QY 121 gaggtccctcgtatatacgaagaaagggtgcagacgtctggtcacaaggttgatc 180
Db 635 gaggtccctcgtatatacgaagaaagggtgcagacgtctggtcacaaggttgatc 694
QY 181 aacactgagtgcagtttaactccagagagttcctcattcgtgtataaagatggcgctg 240
Db 695 aacactgagtgcagtttaactccagagagttcctcattcgtgtataaagatggcgctg 754
QY 241 gcagcccaaaaaaagccatgaagaagccacaagaagtag 282
Db 755 gcagcccaaaaaaagccatgaagaagccacaagaagtag 796
```

## RESULT 6

AAH73234/C  
ID AAH73234 standard; cDNA; 1144 BP.

AC AAH73234;

DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 4508.

DE Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

OS Homo sapiens.

PN WO200142467-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-US33312.

PR 08-DEC-1999; 99US-0169681.

PR 21-DEC-1999; 99US-0171350.

PR 14-MAR-2000; 2000US-0189315.

PR 12-MAY-2000; 2000US-0203791.

PR 09-JUN-2000; 2000US-0210600.

PR 21-JUL-2000; 2000US-0220114.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI: 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer  
and for assessing and detecting compounds for treating the cancer -

PS Claim 1; Page 986; 1051pp; English.

The invention relates to novel genes (AAH68727-AAH73383) associated with cervical cancer with cytostatic activity. The nucleic acids and encoded polypeptides are useful: to assess if a patient is afflicted with cervical cancer or has a pre-malignant condition; to monitor the progression of cervical cancer or a premalignant condition in a patient; and to select and/or assess the efficacy of a compound or therapy for inhibiting cervical cancer in a patient. The nucleic acids may also be useful for gene therapy.

Sequence 1144 BP; 347 A; 257 C; 245 G; 295 T; 0 other;



```
Query Match          100.0%; Score 282; DB 22; Length 1144;
Best Local Similarity 100.0%; Pred. No. 1,9e-80;
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttcacgaagcttgagaagccttgaactctatcatcgacgctctaccagaactacc 60
   |||
DB 436 ATGTTGACCGAGCTGAGAAAGCCCTTGAACTATCATCGACGCTCTACCAAGTACTCC 377
   |||
QY 61 ctgataaagggaattcccatcgctctacagagatgaacctgaagaatattgctgaacc 120
   |||
DB 376 CTGATTAAGGGCAATTTCCATGCCGTCTACAGGATGACCTGAAGAAATTGCTAGACACC 317
   |||
QY 121 gagtgcctcagatataatcagaagaaagggtgcagacgctctgttcaagaagttgatatc 180
   |||
DB 316 GAGTGTCTCCTCAGTATATCATGAGAAAGGGTCCAGACGCTGTTCAAGAAGTTGATATC 257
   |||
QY 181 aacctgattgtgcagtttaacttccaggaagttcctcatcttgttgataaagaatggcggtg 240
   |||
DB 236 AACACTGATGTGAGATTAACTTCACAGAGTTCTCATTTCTGATGAATGAAGTGGCGGTG 197
   |||
QY 241 gcaagcccaaaaaaagccatgaaagaccacaagaagtag 282
   |||
DB 196 GCAGGCCACAAAAAGCCATGAAGAAAGCCACAAGNGTAG 155
   |||

RESULT 7
AAH72611
ID AAH72611 standard; cDNA; 2329 BP.
AC AAH72611;
XX
XX 19-SEP-2001 (first entry)
DE
DE Human cervical cancer marker nucleic acid 3885.
KW
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX MO200142467-A2.
XX
XX 14-JUN-2001.
XX
XX 08-DEC-2000; 2000MO-US33312.
XX
XX 08-DEC-1999; 99US-0169681.
XX
XX 21-DEC-1999; 99US-0171350.
XX
XX 14-MAR-2000; 2000US-0189315.
XX
XX 12-MAY-2000; 2000US-0203791.
XX
XX 09-JUN-2000; 2000US-0210600.
XX
XX 21-JUL-2000; 2000US-0220114.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
XX
XX and for assessing and detecting compounds for treating the cancer -
XX
XX Claim 1; Page 759-760; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX
XX polypeptides are useful: to assess if a patient is afflicted with
XX
XX cervical cancer or has a pre-malignant condition; to monitor the
XX
XX progression of cervical cancer or a premalignant condition in a patient;
XX
XX and to select and/or assess the efficacy of a compound or therapy for
XX
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX
XX useful for gene therapy.
```

```
XX
SQ Sequence 2329 BP; 715 A; 398 C; 509 G; 707 T; 0 other;

Query Match          98.7%; Score 278.4; DB 22; Length 2329;
Best Local Similarity 99.6%; Pred. No. 3,6e-79;
Matches 279; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 gttgacgaagcttgagaagccttgaactctatcatcgacgctctaccagaactacc 62
   |||
DB 51 gttgacgaagcttgagaagccttgaactctatcatcgacgctctaccagaactacc 110
   |||
QY 63 gataaagggaattcccatcgctctacagaggaatgaacctgaagaatlyctagagaccga 122
   |||
DB 111 gataaagggaattcccatcgctctacagaggaatgaacctgaagaatlyctagagaccga 170
   |||
QY 123 gttgcctcagatataatcagaagaaagggtgcagacgctctgttcaagaagttgatatcaa 182
   |||
DB 171 gttgcctcagatataatcagaagaaagggtgcagacgctctgttcaagaagttgatatcaa 230
   |||
QY 183 cactgattgtgcagtttaacttccaggaagttcctcatctctgttgataaagaatggcggtg 242
   |||
DB 231 cactgattgtgcagtttaacttccctgtgagttcctcatctctgttgataaagaatggcggtg 290
   |||
QY 243 agcccaaaaaaagccatgaaagaagccacaagaagtag 282
   |||
DB 291 agcccaaaaaaagccatgaaagaagccacaagaagtag 330
   |||

RESULT 8
AAK52493
ID AAK52493 standard; cDNA; 645 BP.
AC AAK52493;
XX
XX 06-NOV-2001 (first entry)
DE
DE Human polynucleotide SEQ ID NO 2022.
XX
XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
XX
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX
XX nervous system disorder; arthritis; inflammation; ss.
XX
XX Homo sapiens.
OS
XX MO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001MO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX
XX 27-APR-2000; 2000US-0560875.
XX
XX 20-JUN-2000; 2000US-0596075.
XX
XX 19-JUL-2000; 2000US-0620325.
XX
XX 01-SEP-2000; 2000US-0654936.
XX
XX 15-SEP-2000; 2000US-0663561.
XX
XX 20-OCT-2000; 2000US-0693325.
XX
XX 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX
XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
XX
XX P-PSDB; AAM79360.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX
XX useful in diagnosis and gene therapy -
XX
```

XX  
PI  
Schlegel R.  
Needs T.  
Berner A.  
Zhao X.

XX DR WPI; 2001-375006/39.  
XX PT New isolated nucleic acid for diagnosing and treating cervical cancer  
XX PT and for assessing and detecting compounds for treating the cancer -  
XX PS Claim 1; Page 463; 1051pp; English.

XX CC The invention relates to novel genes (AAH68727-AAH73383) associated with  
XX CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
XX CC polypeptides are useful: to assess if a patient is afflicted with  
XX CC cervical cancer or has a pre-malignant condition; to monitor the  
XX CC progression of cervical cancer or a premalignant condition in a patient;  
XX CC and to select and/or assess the efficacy of a compound or therapy for  
XX CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
XX CC useful for gene therapy.

XX SQ Sequence 497 BP; 150 A; 98 C; 118 G; 131 T; 0 other;

Query Match 77.4%; Score 218.4; DB 22; Length 497;  
Best Local Similarity 99.1%; Pred. No. 3.6e-60;  
Matches 230; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 51 caagtaacctgataaagggaattccatgcgctctacaggaatgacctgaagaatt 110  
|||  
DB 36 caggtactccctgataaagggaattccatgcgctctacaggaatgacctgaagaatt 95  
OY 111 gctagaagaccgagtgctccctcagtatatcaggaagggtgcagcgtctgttcaaga 170  
|||||  
DB 96 gctagaagaccgagtgctccctcagtatatcaggaagggtgcagcgtctgttcaaga 155  
OY 171 gttgatatcaaacactgagtgtgcagtttaactccaagagttccctcattctgtgataa 230  
|||||  
DB 156 gttgatatcaaacactgagtgtgcagtttaactccaagagttccctcattctgtgataa 214  
OY 231 gatggcgctgagcagcccaaaaagccatgaagaagcccaagaagtg 282  
|||||  
DB 215 gatggcgctgagcagcccaaaaagccatgaagaagcccaagaagtg 266

RESULT 11  
AAH71815  
ID AAH71815 standard; cDNA; 598 BP.

XX AC AAH71815;

XX DT 19-SEP-2001 (first entry)

XX DE Human cervical cancer marker nucleic acid 3089.

XX KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200142467-A2.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000MO-US33312.

XX PR 08-DEC-1999; 99US-0169681.

XX PR 21-DEC-1999; 99US-0171350.

XX PR 14-MAR-2000; 2000US-0189315.

XX PR 12-MAY-2000; 2000US-0203791.

XX PR 09-JUN-2000; 2000US-0210600.

XX PR 21-JUL-2000; 2000US-0220114.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Deeds J, Berger A, Zhao X;

XX DR WPI; 2001-375006/39.

XX DR New isolated nucleic acid for diagnosing and treating cervical cancer  
XX PT and for assessing and detecting compounds for treating the cancer -  
XX PS Claim 1; Page 616; 1051pp; English.

XX CC The invention relates to novel genes (AAH68727-AAH73383) associated with  
XX CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
XX CC polypeptides are useful: to assess if a patient is afflicted with  
XX CC cervical cancer or has a pre-malignant condition; to monitor the  
XX CC progression of cervical cancer or a premalignant condition in a patient;  
XX CC and to select and/or assess the efficacy of a compound or therapy for  
XX CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
XX CC useful for gene therapy.

XX SQ Sequence 598 BP; 179 A; 113 C; 158 G; 148 T; 0 other;

Query Match 77.4%; Score 218.4; DB 22; Length 598;  
Best Local Similarity 99.1%; Pred. No. 3.9e-60;  
Matches 230; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 51 caagtaacctgataaagggaattccatgcgctctacaggaatgacctgaagaatt 110  
|||  
DB 12 caggtactccctg-taaagggaattccatgcgctctacaggaatgacctgaagaatt 70  
OY 111 gctagaagaccgagtgctccctcagtatatcaggaagggtgcagcgtctgttcaaga 170  
|||||  
DB 71 gctagaagaccgagtgctccctcagtatatcaggaagggtgcagcgtctgttcaaga 130  
OY 171 gttgatatcaaacactgagtgtgcagtttaactccaagagttccctcattctgtgataa 230  
|||||  
DB 131 gttgatatcaaacactgagtgtgcagtttaactccaagagttccctcattctgtgataa 190  
OY 231 gatggcgctgagcagcccaaaaagccatgaagaagcccaagaagtg 282  
|||||  
DB 191 gatggcgctgagcagcccaaaaagccatgaagaagcccaagaagtg 242

RESULT 12  
AAH70613  
ID AAH70613 standard; cDNA; 625 BP.

XX AC AAH70613;

XX DT 19-SEP-2001 (first entry)

XX DE Human cervical cancer marker nucleic acid 1887.

XX KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200142467-A2.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000MO-US33312.

XX PR 08-DEC-1999; 99US-0169681.

XX PR 21-DEC-1999; 99US-0171350.

XX PR 14-MAR-2000; 2000US-0189315.

XX PR 12-MAY-2000; 2000US-0203791.

XX PR 09-JUN-2000; 2000US-0210600.

XX PR 21-JUL-2000; 2000US-0220114.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Deeds J, Berger A, Zhao X;

XX DR WPI; 2001-375006/39.

XX PT New isolated nucleic acid for diagnosing and treating cervical cancer

PT and for assessing and detecting compounds for treating the cancer -  
PS Claim 1; Page 406-407; 1051pp; English.  
XX The invention relates to novel genes (AAH68727-AAH73383) associated with  
CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
CC polypeptides are useful: to assess if a patient is afflicted with  
CC cervical cancer or has a pre-malignant condition; to monitor the  
CC progression of cervical cancer or a premalignant condition in a patient;  
CC and to select and/or assess the efficacy of a compound or therapy for  
CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
CC useful for gene therapy.  
SQ Sequence 625 BP; 184 A; 119 C; 167 G; 155 T; 0 other;

Query Match 77.4%; Score 218.4; DB 22; Length 625;  
Best Local Similarity 99.1%; Pred. No. 3.9e-60;  
Matches 230; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 51 caagtaactccctgataaagggaatttcacatgcgcgtctacaggaatgacctgaagaatt 110  
DB 39 caggtactccctg-taaagggaatttcacatgcgcgtctacaggaatgacctgaagaatt 97  
QY 111 gctagagacgagtgctctcagtatatcaggaagaagggtgcagacgtctgttcaaga 170  
DB 98 gctagagacgagtgctctcagtatatcaggaagaagggtgcagacgtctgttcaaga 157  
QY 171 gttgatatcaacactgtagtgtagttaaacttcacaggaatttcctcattctgtgataa 230  
DB 158 gttgatatcaacactgtagtgtagttaaacttcacaggaatttcctcattctgtgataa 217  
QY 231 gatggggtgtagcagcccaaaaaaagccatgagaagaagccacaagaagtag 282  
DB 218 gatggggtgtagcagcccaaaaaaagccatgagaagaagccacaagaagtag 269

RESULT 13  
AAH71977  
ID AAH71977 standard; cDNA; 381 BP.  
XX  
AC AAH71977;  
XX  
DT 19-SEP-2001 (first entry)  
XX  
DE Human cervical cancer marker nucleic acid 3251.  
XX  
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200142467-A2.  
XX  
PD 14-JUN-2001.  
XX  
PE 08-DEC-2000; 2000WO-US33312.  
XX  
PR 08-DEC-1999; 99US-0169681.  
PR 21-DEC-1999; 99US-0171350.  
PR 14-MAR-2000; 2000US-0189315.  
PR 12-MAY-2000; 2000US-0203791.  
PR 09-JUN-2000; 2000US-0210600.  
PR 21-JUL-2000; 2000US-0220114.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Deeds J, Berger A, Zhao X;  
XX  
DR WPI; 2001-375006/39.  
XX  
PT New isolated nucleic acid for diagnosing and treating cervical cancer  
PT and for assessing and detecting compounds for treating the cancer -  
XX

PS Claim 1; Page 638; 1051pp; English.  
XX The invention relates to novel genes (AAH68727-AAH73383) associated with  
CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
CC polypeptides are useful: to assess if a patient is afflicted with  
CC cervical cancer or has a pre-malignant condition; to monitor the  
CC progression of cervical cancer or a premalignant condition in a patient;  
CC and to select and/or assess the efficacy of a compound or therapy for  
CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
CC useful for gene therapy.  
SQ Sequence 381 BP; 116 A; 78 C; 90 G; 97 T; 0 other;

Query Match 70.6%; Score 199; DB 22; Length 381;  
Best Local Similarity 96.6%; Pred. No. 5.5e-54;  
Matches 225; Conservative 0; Mismatches 5; Indels 3; Gaps 2;  
QY 51 caagtaactccctgataaagggaatttcacatgcgcgtctacaggaatgacctgaagaatt 110  
DB 10 caggtactccctgataaagggaatttcacatgcgcgtctacaggaatgacctgaagaatt 69  
QY 111 gctagagacgagtgctctcagtatatcaggaagaagggtgcagacgtctgttcaaga 170  
DB 70 gctagagacgagtgctctcagtatatcaggaagaagggtgcagacgtctgttcaaga 129  
QY 171 gttgatatcaacactgtagtgtagttaaacttcacaggaatttcctcattctgtgataa 230  
DB 130 gttgatatcaacactgtagtgtagttaaacttcacaggaatttcctcattctgtgataa 187  
QY 231 gatggg-cgtgtagcccaaaaaaagccatgagaagaagccacaagaagtag 282  
DB 188 gatggggtgtagcagcccaaaaaaagccatgagaagaagccacaagaagtag 240

RESULT 14  
AAH69112  
ID AAH69112 standard; cDNA; 512 BP.  
XX  
AC AAH69112;  
XX  
DT 19-SEP-2001 (first entry)  
XX  
DE Human cervical cancer marker nucleic acid 386.  
XX  
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200142467-A2.  
XX  
PD 14-JUN-2001.  
XX  
PE 08-DEC-2000; 2000WO-US33312.  
XX  
PR 08-DEC-1999; 99US-0169681.  
PR 21-DEC-1999; 99US-0171350.  
PR 14-MAR-2000; 2000US-0189315.  
PR 12-MAY-2000; 2000US-0203791.  
PR 09-JUN-2000; 2000US-0210600.  
PR 21-JUL-2000; 2000US-0220114.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Deeds J, Berger A, Zhao X;  
XX  
DR WPI; 2001-375006/39.  
XX  
PT New isolated nucleic acid for diagnosing and treating cervical cancer  
PT and for assessing and detecting compounds for treating the cancer -  
PS Claim 1; Page 170-171; 1051pp; English.  
XX

CC The invention relates to novel genes (AAH68727-AAH73383) associated with  
CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
CC polypeptides are useful: to assess if a patient is afflicted with  
CC cervical cancer or has a pre-malignant condition; to monitor the  
CC progression of cervical cancer or a premalignant condition in a patient;  
CC and to select and/or assess the efficacy of a compound or therapy for  
CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
CC useful for gene therapy.

SQ Sequence 512 BP; 156 A; 103 C; 117 G; 134 T; 2 other;

## Query Match

Best Local Similarity 70.6%; Score 199; DB 22; Length 512;

Matches 225; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 51 caagtactccctgataaagggaattccatgcctctacaggaatgagccgaagaatt 110  
DB 26 caagtactccctgataaagggaattccatgcctctacaggaatgagccgaagaatt 85  
QY 111 gctagagaccgagtgctcctcagatatacaggaaaaagggtgcagacgtctgtccaaga 170  
DB 86 gctagagaccgagtgctcctcagatatacaggaaaaagggtgcagacgtctgtccaaga 145  
QY 171 gttgagatacaacacccgagtgatcaactccaagagaggtccctccttcgtgtgataaa 230  
DB 146 gttgagatacaacacccgagtgatcaactccaagagaggtccctccttcgtgtgataaa 203  
QY 231 gatggg-cgtggcagcccaaaaaagccatgaagaagaagccacaagaagtag 282  
DB 204 gatggg-cgtggcagcccaaaaaagccatgaagaagaagccacaagaagtag 256

## RESULT 15

AAH71248/c  
ID AAH71248 standard; cDNA: 297 BP.

AAH71248;

19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 2522.

Cervical cancer: cytostatic; pre-malignant condition; gene therapy; ss.

Homo sapiens.

MO200142467-A2.

14-JUN-2001.

08-DEC-2000; 2000MO-US33312.

08-DEC-1999; 99US-0169681.

21-DEC-1999; 99US-0171350.

14-MAR-2000; 2000US-0189315.

12-MAY-2000; 2000US-0203791.

09-JUN-2000; 2000US-0210600.

21-JUL-2000; 2000US-0220114.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Deeds J, Berger A, Zhao X;

WPI: 2001-375006/39.

New isolated nucleic acid for diagnosing and treating cervical cancer  
and for assessing and detecting compounds for treating the cancer -

Claim 1; Page 525; 1051pp; English.

The invention relates to novel genes (AAH68727-AAH73383) associated with  
cervical cancer with cytostatic activity. The nucleic acids and encoded

CC polypeptides are useful: to assess if a patient is afflicted with  
CC cervical cancer or has a pre-malignant condition; to monitor the  
CC progression of cervical cancer or a premalignant condition in a patient;  
CC and to select and/or assess the efficacy of a compound or therapy for  
CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
CC useful for gene therapy.

SQ Sequence 297 BP; 66 A; 77 C; 66 G; 88 T; 0 other;

## Query Match

Best Local Similarity 67.7%; Score 191; DB 22; Length 297;

Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 ggaatgacctgagaagaattgctagagaccgagtgctcctcagatatacaggaagaaggtc 151  
DB 297 ggaatgacctgagaagaattgctagagaccgagtgctcctcagatatacaggaagaaggtc 238  
QY 152 cgaagctcgtgtcaaaagattgatatcaaacactgaatggtgcaagtaacttcaggaagt 211  
DB 237 cgaagctcgtgtcaaaagattgatatcaaacactgaatggtgcaagtaacttcaggaagt 178  
QY 212 tctcatctgtgtgataaagatgagcgtgagcccaaaaaaagccatgaagaagcc 271  
DB 177 tctcatctgtgtgataaagatgagcgtgagcccaaaaaaagccatgaagaagcc 118  
QY 272 acaagaagtag 282  
DB 117 ACAAGAGTAG 107

Search completed: September 10, 2002, 06:49:27  
Job time: 5203 sec

.....

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2002, 04:17:24 ; Search time 2463.79 Seconds  
(without alignments)  
1544.832 Million cell updates/sec

Title: US-09-806-382A-1

Perfect score: 282  
Sequence: 1 atgttgaccgagctcgagaa.....aagaagccacaagagtag 282

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estinu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_estc2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_dln:\*  
16: em\_gss\_vrl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	282	100.0	324	9	AA086471 zn62b01.s
C 2	282	100.0	325	9	AA086471 zn62b01.s
C 3	282	100.0	337	9	AA584358
C 4	282	100.0	361	10	F24145
C 5	282	100.0	369	9	AA086471 zn62b01.s
C 6	282	100.0	376	9	AA086471 zn62b01.s
C 7	282	100.0	376	9	AA086471 zn62b01.s
C 8	282	100.0	377	9	AA086471 zn62b01.s
C 9	282	100.0	380	9	AA086471 zn62b01.s
C 10	282	100.0	384	9	AA086471 zn62b01.s
C 11	282	100.0	384	10	BF478287
C 12	282	100.0	386	9	AA086471 zn62b01.s
C 13	282	100.0	388	9	AA086471 zn62b01.s
C 14	282	100.0	389	9	AA086471 zn62b01.s
C 15	282	100.0	390	10	W79564
C 16	282	100.0	391	9	AA086471 zn62b01.s
C 17	282	100.0	394	9	AA086471 zn62b01.s

18	282	100.0	396	9	AA086471 zn62b01.s
19	282	100.0	396	9	AA086471 zn62b01.s
20	282	100.0	397	9	AA086471 zn62b01.s
21	282	100.0	398	9	AA086471 zn62b01.s
22	282	100.0	399	9	AA086471 zn62b01.s
C 23	282	100.0	400	9	AA086471 zn62b01.s
C 24	282	100.0	400	10	BF478287
C 25	282	100.0	404	9	AA086471 zn62b01.s
C 26	282	100.0	406	9	AA086471 zn62b01.s
C 27	282	100.0	406	9	AA086471 zn62b01.s
C 28	282	100.0	407	9	AA086471 zn62b01.s
C 29	282	100.0	407	10	BF478287
C 30	282	100.0	410	9	AA086471 zn62b01.s
C 31	282	100.0	410	10	BF478287
C 32	282	100.0	414	9	AA086471 zn62b01.s
C 33	282	100.0	416	9	AA086471 zn62b01.s
C 34	282	100.0	416	9	AA086471 zn62b01.s
C 35	282	100.0	418	9	AA086471 zn62b01.s
C 36	282	100.0	419	9	AA086471 zn62b01.s
C 37	282	100.0	419	10	BF478287
C 38	282	100.0	420	9	AA086471 zn62b01.s
C 39	282	100.0	421	9	AA086471 zn62b01.s
C 40	282	100.0	422	9	AA086471 zn62b01.s
C 41	282	100.0	423	9	AA086471 zn62b01.s
C 42	282	100.0	424	9	AA086471 zn62b01.s
C 43	282	100.0	427	9	AA086471 zn62b01.s
C 44	282	100.0	428	9	AA086471 zn62b01.s
C 45	282	100.0	428	9	AA086471 zn62b01.s

## ALIGNMENTS

RESULT 1  
LOCUS AA086471/c  
DEFINITION zn62b01.s1 StrataGene muscle 937209 Homo sapiens cDNA clone IMAGE:562729 3' similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);  
RNA sequence.  
AA086471.1 GI:1629088  
EST.

ACCESSION AA086471.1 GI:1629088  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chappell, B., Chisoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W., Hawkins, B., Morris, M., Parsons, J., Prange, C., Rife, L., Rohlfs, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

TITLE  
JOURNAL MEDLINE  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
Image Consortium (info@image.llnl.gov) for further information.  
Insert Length: 450 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham.

## FEATURES

Location/Qualifiers  
1..324  
/db\_xref="taxon:9606"  
/db\_xref="taxon:9606"  
/db\_xref="taxon:9606"  
/clone="IMAGE:562729"  
/clone\_lib="Stratagene muscle 937209"

/tissue\_type="muscle"  
/dev\_stage="adult"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: skeletal muscle; Vector: pBluescript SK-  
Site: 1: EcorI; Site 2: XhoI; Cloned unidirectionally.  
Primer: Oligo dt. Skeletal muscle from patient with  
malignant hyperthermia. Average insert size: 1.0 kb;  
uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGCAG  
3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'"

BASE COUNT 70 a 84 c 70 g 97 t 3 others  
ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 324;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgaccgagctggagaagccttgaactctatcatcagcgtctaccagaagtacc 60  
|||||  
DB 321 ATGTTGACCGAGCTGGAGAAAGCCTTGACTCTATCATCGACGCTACCAAGTACTCC 262  
QY 61 ctgataaagggaattccatccgctctcaaggatgacctgaagaattgctagagacc 120  
|||||  
DB 261 CTGATTAAGGGGAATTTCCATGCCCTCTACAGGATGACCTGAAGAATTGCTAGAGACC 202  
QY 121 gagtgcctcagtatatcaggaagagggttcagagcgtctggttcaagaagttgatatc 180  
|||||  
DB 201 GAGTGTCTCAGTATATCAGAAAAAGGTGCAAGACCTGTGTTCAAGAGTTGATATC 142  
QY 181 aacctgatggtgagtaacttccaggagttctctcatctctgtgtataaagatggcgctg 240  
DB 141 AACACTGATGTGTCAGTTAACTTCACAGAGTTCTCTCATTCGTGATTAAGATGGCGCTG 82  
QY 241 gcagcccaaaaaagccatgaagaagccacaaagaagtag 282  
|||||  
DB 81 GCAGCCCAAAAAAGCCATGAAGAAAGCCACAAGAGTAG 40

RESULT 2  
AM238704 325 bp mRNA linear EST 13-DEC-1999  
LOCUS  
DEFINITION xp306i2.x1 NCI-CGAP\_HN10 Homo sapiens cDNA clone IMAGE:2741902 3'  
similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);, mRNA sequence.  
AM238704  
ACCESSION AM238704.1 GI:6571093  
VERSION  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 325)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: Edward Shillito Ph.D., Silvio Gutkind Ph.D.,  
Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

FEATURES  
SOURCE  
Possible reversed clone: polyA not found  
Seq primer: -40UP from Gidco  
High quality sequence stop: 301.  
Location/Qualifiers  
1..325  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

/clone="IMAGE:2741902"  
/clone\_lib="NCI CGAP\_HN10"  
/tissue\_type="carcinoma in situ from retroocular trigone"  
/lab\_host="DH10B"  
/note="Vector: pAMP10; CDNA made by Oligo-dT priming.  
Non-directionally cloned into the UDG sites of pAMP10.  
Size-selected on agarose gel, average insert size 500 bp.  
Primary library; non-amplified. CDNA Library  
Preparation: David B. Krizman, Ph.D (NCI). Reference:  
Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 96 a 72 c 89 g 68 t  
ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgaccgagctggagaagccttgaactctatcatcagcgtctaccagaagtacc 60  
|||||  
DB 16 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATGACGCTACCAAGTACTCC 75  
QY 61 ctgataaagggaattccatccgctctcaaggatgacctgaagaattgctagagacc 120  
|||||  
DB 76 CTGATTAAGGGGAATTTCCATGCCCTCTACAGGATGACCTGAAGAATTGCTAGAGACC 135  
QY 121 gagtgcctcagtatatcaggaagagggttcagagcgtctggttcaagaagttgatatc 180  
|||||  
DB 136 GACTGTCTCAGTATATCAGAAAAAGGTGCAAGCTGTGTTCAAGAGTTGATATC 195  
QY 181 aacctgatggtgagtaacttccaggagttctctcatctctgtgtataaagatggcgctg 240  
DB 196 AACACTGATGTGTCAGTTAACTTCACAGAGTTCTCTCATTCGTGATTAAGATGGCGCTG 255  
QY 241 gcagcccaaaaaagccatgaagaagccacaaagaagtag 282  
|||||  
DB 256 GCAGCCCAAAAAAGCCATGAAGAAAGCCACAAGAGTAG 297

RESULT 3  
AA584358/c 337 bp mRNA linear EST 26-SEP-1997  
LOCUS  
DEFINITION AA584358.s1 NCI-CGAP\_Lari Homo sapiens cDNA clone IMAGE:1089515 3'  
similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);, mRNA sequence.  
AA584358  
ACCESSION AA584358.1 GI:2368967  
VERSION  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 337)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Stratagene, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 534 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amerisham.

FEATURES  
SOURCE  
Location/Qualifiers  
1..337  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1089515"



/clone\_1lb="NCI\_CGAP\_Lar1"  
/issue\_type="larynx"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: larynx; Vector: Bluescript SK-; Site: 1;  
ECORI: Site: 2; XhoI: Cloned unidirectionally. Primer:  
0.1lgo dt. larynx. 5' adaptor sequence. 5' GAATTCGGCAGC  
3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'  
Average insert size: 0.9 kb."  
BASE COUNT 73 a 89 c 76 g 99 t  
ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 337;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgttaccgagcttgagaagccttgaaactatcatcagcgtctacacaagtaactcc 60  
DB 323 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCATCATGACGTCTACACAAAGTACTCC 264  
QY 61 ctgataaagggaattccatgcgcgtctacaggagatgacctaagaataatgtctagaagcc 120  
DB 263 CTGATTAAGGGGAATTTCCATGCGCGTCTACAGGAGATGACCTGAAGAAATTTGCTAGAGACC 204  
QY 121 gaftgtcctaagatatacaggaagaggtgcagacgtctggttcaagaagttgataatc 180  
DB 203 GAGTGTCTCAGTATATACAGGAAAGGGTGCAGACGTCTGTTCAAGAGCTTGATATAC 144  
QY 181 aacactgagtgtcgaattactccagagagttccattctgtgataaagatggagctg 240  
DB 143 AACCTGATGTCTCAGTATTAATCTTCAGAGAGTTCCTCATTTCTGTATTAAGATGGCGGTG 84  
QY 241 gcagcccaaaaaaaagccatgagaagaaagccacaagaagtag 282  
DB 83 GCAGCCCAAAAAAAGCCATGAAGAAAGCCACAAAGAGTAG 42

RESULT 4  
F24145 361 bp mRNA linear EST 13-MAY-1999  
LOCUS HSPD10203 HM3 Homo sapiens cDNA clone s4000003611, mRNA sequence.  
DEFINITION F24145  
ACCESSION F24145.1 GI:4809771  
VERSION EST.  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 361)  
AUTHORS Lanfranchi, G., Muraro, T., Caldera, F., Pachioni, B., Pallavicini, A.,  
Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.  
TITLE Identification of 4370 expressed sequence tags from a  
3'-end-specific cDNA library of human skeletal muscle by DNA  
sequencing and filter hybridization  
JOURNAL Genome Res. 6 (1), 35-42 (1996)  
MEDLINE 96276048  
COMMENT Contact: Valle G.  
CIRIBI Biotechnology Centre  
University of Padua  
Via Trieste 75, 35121 Padua, Italy  
ABI Chromatograms and other information are available on WWW at  
http://grup.bio.unipd.it.  
location/Qualifiers  
1. 361  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="s4000003611"  
/clone\_1lb="HM3"  
/sex="female"  
/issue\_type="pectoral muscle (after mastectomy)"  
/note="Vector: pCDNAII (Invitrogen); Site: 1: BstXI;  
Site: 2: NotI. The library was constructed by G.  
Lanfranchi. This library is not subtracted nor normalized.

The first strand cDNA was primed with a biotinylated  
0.1lgo-dT-NotI primer  
(5'-biotin-AACCGGCTCGACGCGCCGCTTTTCTTTTCTTTT-3'). The  
ds cDNA was sonicated and size-selected in the range  
350-550 bp. The 3' specific fragments were selected by  
streptavidin coated magnetic beads, ligated to  
non-palindromic BstXI adaptors, NotI digested and  
directionally cloned into BstXI-NotI cut pCDNAII vector."  
BASE COUNT 106 a 80 c 100 g 75 t  
ORIGIN

Query Match 100.0%; Score 282; DB 10; Length 361;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgttaccgagcttgagaagccttgaaactatcatcagcgtctacacaagtaactcc 60  
DB 32 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCATCATGACGTCTACACAAAGTACTCC 91  
QY 61 ctgataaagggaattccatgcgcgtctacaggagatgacctaagaataatgtctagaagcc 120  
DB 92 CTGATTAAGGGGAATTTCCATGCGCGTCTACAGGAGATGACCTGAAGAAATTTGCTAGAGACC 151  
QY 121 gaftgtcctaagatatacaggaagaggtgcagacgtctggttcaagaagttgataatc 180  
DB 152 GAGTGTCTCAGTATATACAGGAAAGGGTGCAGACGTCTGTTCAAGAGCTTGATATAC 211  
QY 181 aacactgagtgtcgaattactccagagagttccattctgtgataaagatggagctg 240  
DB 212 AACCTGATGTCTCAGTATTAATCTTCAGAGAGTTCCTCATTTCTGTATTAAGATGGCGGTG 271  
QY 241 gcagcccaaaaaaaagccatgagaagaaagccacaagaagtag 282  
DB 272 GCAGCCCAAAAAAAGCCATGAAGAAAGCCACAAAGAGTAG 313

RESULT 5  
AW265292 369 bp mRNA linear EST 28-DEC-1999  
LOCUS xq49h08.x1 NCI\_CGAP\_HN7 Homo sapiens cDNA clone IMAGE:2754015 3'  
DEFINITION similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);, mRNA sequence.  
ACCESSION AW265292  
VERSION AW265292.1 GI:6642108  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 369)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,  
Childanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.  
CDNA Library Preparation: David B. Krieman, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bdrp/image/image.html

FEATURES  
source  
Possible reversed clone: polyT not found  
Seq primer: -40UP from Gibco.  
location/Qualifiers  
1. 369  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2754015"  
/clone\_1lb="NCI\_CGAP\_HN7"

/tissue\_type="normal squamous epithelium, floor of mouth"  
/lab\_host="DH10B"  
/note="Vector: PAMP10; cDNA made by oligo-dT priming.  
/non-directionally cloned into the UDG sites of PAMP10.  
Size-selected on agarose gel, average insert size 500 bp.  
Primary library; non-amplified. cDNA library  
Preparation: David B. Krizman, Ph.D. (NCI). Reference:  
Krizman et al. (1996) Cancer Research 56:5380-5383."  
BASE COUNT 111 a 84 c 96 g 78 t  
ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 369;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttaccagctggaagaagccttgaactctatcatcagcgtctacacaagtactcc 60  
|||||  
DB 20 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCATCATGAGCTGTACCAAGTACTCC 79  
|||||  
QY 61 ctgataaagggaattccatgcgcgtctacaggatgacctgaagaattgctagagacc 120  
|||||  
DB 80 CTGATTAAGGGGAATTCCATGCCGTCTACAGGATGACCTGAAGAATTGCTAGAGACC 139  
|||||  
QY 121 gagtctcctcagtatatcagaaaagggtgcagacgtctggttcaaaagattgagatc 180  
|||||  
DB 140 GAGGTCTCTAGTATATCAGAAAAGGGTGCAGACGTCTGTTCAAAAGTTGGATATC 199  
|||||  
QY 181 aacactgatgtygcagttactccagaggttctcattctggtgataaagtggcggtg 240  
|||||  
DB 200 AACACTGATGGTGAGTTACTTCCAGGAGTTCCTCATCTGTGATTAAGATGGCGTG 259  
|||||

QY 241 gcaagcccaaaaaagccatgaaagaagccacaagaagtag 282  
|||||  
DB 260 GCAGCCCACAAAAAAGCCATGAAGAGCCACAAAGAGTAG 301  
|||||

RESULT 6  
BE139192/c 369 bp mRNA linear EST 21-JUN-2000  
LOCUS BE139192  
DEFINITION xp38610.x2 NCI\_CGAP\_Lu28 Homo sapiens cDNA clone IMAGE:2752915 3'  
similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);, mRNA sequence.  
ACCESSION BE139192  
VERSION BE139192.1 GI:8601615  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 369)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life  
Technologies, Inc. cDNA Library Arrayed by: Christina Prange, The  
Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:  
[info@image.lnl.gov](mailto:info@image.lnl.gov)  
Seq primer: -40up from Gibco.  
Location/Qualifiers  
1. 369  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2752915"  
/clone\_lib="NCI\_CGAP\_Lu28"  
/tissue\_type="NCI\_CGAP\_Lu28"  
/tissue\_type="two pooled squamous cell carcinomas"  
/lab\_host="DH10B"

FEATURES  
source

/note="Organ: lung; Vector: PCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies."  
BASE COUNT 79 a 95 c 83 g 112 t  
ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 369;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttaccagctggaagaagccttgaactctatcatcagcgtctacacaagtactcc 60  
|||||  
DB 352 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCATCATGAGCTGTACCAAGTACTCC 293  
|||||  
QY 61 ctgataaagggaattccatgcgcgtctacaggatgacctgaagaattgctagagacc 120  
|||||  
DB 292 CTGATTAAGGGGAATTCCATGCCGTCTACAGGATGACCTGAAGAATTGCTAGAGACC 233  
|||||  
QY 121 gagtctcctcagtatatcagaaaagggtgcagacgtctggttcaaaagattgagatc 180  
|||||  
DB 232 GAGTCTCTCACTATATCAGAAAAGGGTGCAGACGTCTGTTCAAAAGTTGGATATC 173  
|||||  
QY 181 aacactgatgtygcagttactccagaggttctcattctggtgataaagtggcggtg 240  
|||||  
DB 172 AACACTGATGGTGAGTTACTTCCAGGAGTTCCTCATCTGTGATTAAGATGGCGTG 113  
|||||

QY 241 gcaagcccaaaaaagccatgaaagaagccacaagaagtag 282  
|||||  
DB 112 GCAGCCCACAAAAAAGCCATGAAGAGCCACAAAGAGTAG 71  
|||||

RESULT 7  
AW238650/c 376 bp mRNA linear EST 13-DEC-1999  
LOCUS AW238650  
DEFINITION xp29f02.x1 NCI\_CGAP\_HN10 Homo sapiens cDNA clone IMAGE:2741787 3'  
similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);, mRNA sequence.  
ACCESSION AW238650  
VERSION AW238650.1 GI:6571039  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 376)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,  
Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
[www.bio.lnl.gov/bbrp/image/image.html](http://www.bio.lnl.gov/bbrp/image/image.html)  
Possible reversed clone: polyT not found  
Seq primer: -40up from Gibco  
High quality sequence stop: 326.  
Location/Qualifiers  
1. 376  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2741787"  
/clone\_lib="NCI\_CGAP\_HN10"  
/tissue\_type="carcinoma in situ from retromolar trigone"  
/lab\_host="DH10B"  
/note="Vector: PAMP10; cDNA made by oligo-dT priming.  
Non-directionally cloned into the UDG sites of PAMP10.

FEATURES  
source

Size-selected on agarose gel, average insert size 500 bp.  
 Primary library; non-amplified. CDNA Library  
 Preparation: David B. Krizman, Ph.D (NCI). Reference:  
 Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 77 a 92 c 81 g 126 t

ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 376;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttgaccgagctgagagaagccttgactctatcaltcagcgtctaccacaagtactcc 60  
 |||||||  
 Db 376 ATGTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATCGAGCTTACCAAGTACTCC 317  
 QY 61 ctgataaagggaatttcacatgcggtctacaggagatgaccctgaagaatttctagagacc 120  
 |||||||  
 Db 316 CTGATMAAGGGGAATTTCCATGCCCTCTACAGGATGACCTGAAGAAATTTCTAGAGACC 257  
 QY 121 gagtgctcagatatacaggaagaaaggctgacagcgtctgttcaagaagtgtgatalc 180  
 |||||||  
 Db 256 GAGTGTCTCTCACTATATATGAGAAAAAGGCTGACAGACGCTGTCTCAAGAGATTGGATATC 197  
 QY 181 aacactgatagtgtcagtaacttccaggaagtctcattctctgtgtatgaagtggcgctg 240  
 |||||||  
 Db 196 AACACTGATGGTGGCGATTAACTTCAGAGATTCTCTCATTTCTGGATTAAGATGGGCGCTG 137  
 QY 241 gcagcccaaaaaaagccatgtaagaagccacaaagatgag 282  
 |||||||  
 Db 136 GCAGCCACAAAAAAGCCATGAAGAAAGCCACAAAGAGTAG 95

RESULT 8  
 LOCUS Aa112727 377 bp mRNA linear EST 01-AUG-1997  
 DEFINITION zn62b01.x1 Stratiagene muscle 937209 Homo sapiens cDNA clone  
 IMAGE:562729 5' similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);,  
 mRNA sequence.

ACCESSION Aa112727  
 VERSION Aa112727.1 GI:1663860  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 377)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
 Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,  
 M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,  
 B., Morris, M., Parsons, J., Rife, C., Rifkin, L., Rohlfing, T.,  
 Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E.,  
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Mairra, M.  
 Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)

TITLE  
 JOURNAL MEDLINE  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estewatson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Putative full length read  
 The vector to vector length is 378  
 Insert Length: 450 Std Error: 0.00  
 Seq primer: -28M13 rev2 from Amersham  
 High quality sequence stop: 356.

FEATURES  
 source  
 1..377  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="GDB:456019"

/db\_xref="taxon:9606"  
 /clone="IMAGE:562729"  
 /clone\_lib="Stratiagene muscle 937209"  
 /tissue\_type="muscle"  
 /dev\_stage="adult"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /note="Organ: skeletal muscle; Vector: p Bluescript SK-;  
 Site\_1: EcoRI; Site\_2: XhoI; Cloned unidirectionally.  
 Primer: Oligo dt. Skeletal muscle from patient with  
 malignant hyperthermia. Average insert size: 1.0 kb;  
 Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG  
 3'-3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

BASE COUNT 113 a 83 c 100 g 81 t

ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttgaccgagctgagagaagccttgactctatcaltcagcgtctaccacaagtactcc 60  
 |||||||  
 Db 23 ATGTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATCGAGCTTACCAAGTACTCC 82  
 QY 61 ctgataaagggaatttcacatgcggtctacaggagatgaccctgaagaatttctagagacc 120  
 |||||||  
 Db 83 CTGATMAAGGGGAATTTCCATGCCCTCTACAGGATGACCTGAAGAAATTTCTAGAGACC 142  
 QY 121 gagtgctcagatatacaggaagaaaggctgacagcgtctgttcaagaagtgtgatalc 180  
 |||||||  
 Db 143 GAGTGTCTCTCACTATATATGAGAAAAAGGCTGACAGACGCTGTCTCAAGAGATTGGATATC 202  
 QY 181 aacactgatagtgtcagtaacttccaggaagtctcattctctgtgtatgaagtggcgctg 240  
 |||||||  
 Db 203 AACACTGATGGTGGCGATTAACTTCAGAGATTCTCTCATTTCTGGATTAAGATGGGCGCTG 262  
 QY 241 gcagcccaaaaaaagccatgtaagaagccacaaagatgag 282  
 |||||||  
 Db 263 GCAGCCACAAAAAAGCCATGAAGAAAGCCACAAAGAGTAG 304

RESULT 9  
 LOCUS AW265343 380 bp mRNA linear EST 28-DEC-1999  
 DEFINITION xg50h04.x1 NCI-CGAP\_HN7 Homo sapiens cDNA clone IMAGE:2754103 3'  
 similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);,  
 mRNA sequence.

ACCESSION AW265343  
 VERSION AW265343.1 GI:6642159  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 380)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)

JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,  
 Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.  
 CDNA Library Preparation: David B. Krizman, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www-bio.llnl.gov/dbp/image/image.html

FEATURES  
 source  
 1..377  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="GDB:456019"

Possible reversed clone: polyT not found  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 370.

```
source
1. .380
/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:2754103"
/clone_1ib="NCI_CGAP_HN7"
/tissue_type="normal squamous epithelium, floor of mouth"
/lab_host="DH10B"
/Note="Vector: PAMPI0; cDNA made by oligo-dt priming.
Non-directionally cloned into the UDG sites of PAMPI0.
Size-selected on agarose gel, average insert size 500 bp.
Primary library; non-amplified. cDNA Library
Preparation: David B. Krizman, Ph.D (NCI). Reference:
Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT      109 a      85 c      104 g      82 t
ORIGIN

Query Match      100.0%; Score 282; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.8e-66;
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttaccgagctggagaagccttgaacctatctatcgagctctaccagaattacc 60
Db 48 ATGTTGACCGAGCTGGAAGAGCCCTTGAACTCATCATCGACGCTACCAACAGTACTCC 107
QY 61 ctgataaagggaattccatgcccgtctacaggatgacctgaagaattgctagagacc 120
Db 108 CTGATTAAGGGGAATTTCCATGCCGCTCAACAGGATGACCTGAAGAAATTGCTAGAGACC 167
QY 121 gaggtccctcgtatatacagaagaagggtgcagagctcgtgttcaagaagtggatatac 180
Db 168 GAGGTCTCTCGTATATCAGGAAAAAGGTCACACGCTCGTTCAAGAGATTGATATC 227
QY 181 aacactgatgtgcagtttaactccaggagttctctatctgtgtgataagaatggggtg 240
Db 228 AACCTGATGTGTCAGTTACTTCCAGAGTTCTCTCATTTGTTGATTAAGATGGCGTG 287
QY 241 gcagcccaacaataaagccatgaagaagccacaaagaagtat 282
Db 288 GCAGCCCAACAATAAAGCCATGAAGAAAGCCACAAAGATGAG 329

RESULT 10
AM238239      384 bp      mRNA      linear      EST 13-DEC-1999
LOCUS      AM238239
DEFINITION      xp230405.x1 NCI_CGAP_HN10 Homo sapiens cDNA clone IMAGE:2741193 3'
ACCESSION      AM238239
VERSION      AM238239.1 GI:6570711
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 384)
NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Edward J. Shillitoe Ph.D., Silvio Gutkind Ph.D.,
Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbfp/image/image.html

Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 367.
```

```
FEATURES
source
Location/Qualifiers
1. .384
/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:2741193"
/clone_1ib="NCI_CGAP_HN10"
/tissue_type="carcinoma in situ from retromolar trigone"
/lab_host="DH10B"
/Note="Vector: PAMPI0; cDNA made by oligo-dt priming.
Non-directionally cloned into the UDG sites of PAMPI0.
Size-selected on agarose gel, average insert size 500 bp.
Primary library; non-amplified. cDNA Library
Preparation: David B. Krizman, Ph.D (NCI). Reference:
Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT      113 a      85 c      105 g      81 t
ORIGIN

Query Match      100.0%; Score 282; DB 9; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.8e-66;
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttaccgagctggagaagccttgaacctatctatcgagctctaccagaattacc 60
Db 46 ATGTTGACCGAGCTGGAAGAGCCCTTGAACTCATCATCGAGCTTACCAAGATCTCC 105
QY 61 ctgataaagggaattccatgcccgtctacaggatgacctgaagaattgctagagacc 120
Db 106 CTGATTAAGGGGAATTTCCATGCCGCTCAACAGGATGACCTGAAGAAATTGCTAGAGACC 165
QY 121 gaggtccctcgtatatacagaagaagggtgcagagctcgtgttcaagaagtggatatac 180
Db 166 GAGGTCTCTCGTATATCAGGAAAAAGGTCACACGCTCGTTCAAGAGATTGATATC 225
QY 181 aacactgatgtgcagtttaactccaggagttctctatctgtgtgataagaatggggtg 240
Db 226 AACCTGATGTGTCAGTTACTTCCAGAGTTCTCTCATTTGTTGATTAAGATGGCGCTG 285
QY 241 gcagcccaacaataaagccatgaagaagccacaaagaagtat 282
Db 286 GCAGCCCAACAATAAAGCCATGAAGAAAGCCACAAAGATGAG 327

RESULT 11
BF478287/c      384 bp      mRNA      linear      EST 05-DEC-2000
LOCUS      BF478287/c
DEFINITION      nae46908.x1 NCI_CGAP_HN16 Homo sapiens cDNA clone IMAGE:4115127 3'
ACCESSION      BF478287
VERSION      BF478287.1 GI:11549114
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 384)
NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Edward J. Shillitoe, Ph.D., Chidchanok
Leethanakul
cDNA Library Preparation: D. Krizman & C. Leethanakul Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .384
```

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone_id="NCI_CGAP_HN16"  
/issue_type="moderate to poorly differentiated invasive  
carcinoma of the retromolar trigone"  
/lab_host="DH10B"  
/note="Organ: mouth; Vector: PAMP10; mRNA made from  
carcinoma, cDNA made by oligo-dT priming.  
Non-directionally cloned into UDG sites. Size-selected on  
agarose gel, average insert size 500 bp. Primary library.  
cDNA Library Preparation: David B. Krizman, Ph.D."
```

BASE COUNT 80 a 97 c 86 g 121 t

ORIGIN

Query Match 100.0%; Score 282; DB 10; Length 384;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttgaccgagcttgagagaagccttgaaactatcatcgcgcgtctacccaagtactcc 60  
|||  
DB 367 AGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATCGACGTCTACACAGTACTCC 308  
|||  
QY 61 ctgataaagggaattccatgcgctctacaggagatgacctgaagaattgctagagacc 120  
|||  
DB 307 CTGATTAAGGGGGAATTTCCATGCCCTTACAGGATGACCTGAAGAAATTCCTAGAGACC 248  
|||  
QY 121 gagtgtcctcagatatacagaagaagggtcagacgctcgtctcaagaagttgataatc 180  
|||  
DB 247 GAGTGTCTCCACTATATCATGAGAAAAAGGTTGACGACGCTGTTCAAGAGATTGGATATC 188  
|||  
QY 181 aacactgagtgtgcagtaacttccaggagttccctcattctggtgataaagtggcgctg 240  
|||  
DB 187 AACACTGATGATGTCAGTTAACTTCACAGAGTTCTCTATTCGTGATTAAGATGGCGCTG 128  
|||  
QY 241 gcaagcccaaaaaaagccatgaagaagccacaagaagtag 282  
|||  
DB 127 GCAGCCCAAAAAAGCCATGAAAGCAAGCAAAAGAGTAG 86  
|||

RESULT 12  
AM238329 386 bp mRNA linear EST 13-DEC-1999  
LOCUS xp21a08.x1 NCI\_CGAP\_HN10 Homo sapiens cDNA clone IMAGE:2740982 3'  
DEFINITION similar to gb:A12027\_cds1 CALGRANDLIN A (HUMAN);, mRNA sequence.  
ACCESSION AM238329  
VERSION AM238329.1 GI:6570646  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 386)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,  
Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.lnl.gov/dbtrp/image/image.html](http://www.bio.lnl.gov/dbtrp/image/image.html)  
Seq primer: -400P from Glibco.  
Location/Qualifiers  
1. 386  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_id="NCI\_CGAP\_HN16"

```
/clone="IMAGE:2740982"  
/clone_id="NCI_CGAP_HN10"  
/issue_type="carcinoma in situ from retromolar trigone"  
/lab_host="DH10B"  
/note="Vector: PAMP10; cDNA made by oligo-dT priming.  
Non-directionally cloned into the UDG sites of PAMP10.  
Size-selected on agarose gel, average insert size 500 bp.  
Primary library; non-amplified. cDNA Library  
Preparation: David B. Krizman, Ph.D (NCI). Reference:  
Krizman et al. (1996) Cancer Research 56:5380-5383."
```

BASE COUNT 122 a 86 c 99 g 79 t

ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 386;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttgaccgagcttgagagaagccttgaaactatcatcgcgcgtctacccaagtactcc 60  
|||  
DB 15 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATCGACGTCTACACAGTACTCC 74  
|||  
QY 61 ctgataaagggaattccatgcgctctacaggagatgacctgaagaattgctagagacc 120  
|||  
DB 75 CTGATTAAGGGGGAATTTCCATGCCCTTACAGGATGACCTGAAGAAATTCCTAGAGACC 134  
|||  
QY 121 gagtgtcctcagatatacagaagaagggtcagacgctcgtctcaagaagttgataatc 180  
|||  
DB 135 GAGTGTCTCCACTATATCATGAGAAAAAGGTTGACGACGCTGTTCAAGAGATTGGATATC 194  
|||  
QY 181 aacactgagtgtgcagtaacttccaggagttccctcattctggtgataaagtggcgctg 240  
|||  
DB 195 AACACTGATGATGTCAGTTAACTTCACAGAGTTCTCTATTCGTGATTAAGATGGCGCTG 254  
|||  
QY 241 gcaagcccaaaaaaagccatgaagaagccacaagaagtag 282  
|||  
DB 255 GCAGCCCAAAAAAGCCATGAAAGCAAGCAAAAGAGTAG 296  
|||

RESULT 13  
AM080385 388 bp mRNA linear EST 14-OCT-1999  
LOCUS x641e08.x1 NCI\_CGAP\_E802 Homo sapiens cDNA clone IMAGE:2610470 3'  
DEFINITION similar to gb:A12027\_cds1 CALGRANDLIN A (HUMAN);, mRNA sequence.  
ACCESSION AM080385  
VERSION AM080385.1 GI:6035537  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 388)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip  
Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.lnl.gov/dbtrp/image/image.html](http://www.bio.lnl.gov/dbtrp/image/image.html)  
Seq primer: -400P from Glibco  
High quality sequence stop: 373.  
Location/Qualifiers  
1. 388  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2610470"

/clone\_lib="NCI-CGAP\_Eso2"  
/tissue\_type="squamous cell carcinoma"  
/lab\_host="DH10B"  
/note="Organ: esophagus; Vector: PCMV-SPORT6; Site\_1: Salt  
; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.1 kb. Life Technologies catalog  
#: 11502-010"

BASE COUNT 84 a 102 c 86 g 116 t  
ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 388;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgttgacgcagctgagaaagccttgaaactatcatcgacgcttaccacaagctacc 60  
DB 352 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTTATCATCGACGTACCAAGTACTCC 293

OY 61 ctgataaagggaattccatgcgctctacagagatgacctgaagaattgctagaacc 120  
DB 292 CTGATTAAGGGGAATTTCCATCGCTACAGGATGACGTGAAGAATTTGCTAGAGACC 233

OY 121 gagttctcagatatacaggaagaggtgcagacgtctggttcaaaagattgatac 180  
DB 232 GAGTGTCTCACTATATACGAGAAAGGTCGACACGTCTGTTCAAGAGTTGATATC 173

OY 181 aacactgattgtcagtttaacttccagagattccctcatcttggtatagaatggcggt 240  
DB 172 AACACTGATGTGTGCACTTACTTCAGAGATTCCTCATTTGTTGATTAAGTGGCGGTG 113

OY 241 gcagcccaaaaaaagccatgaagaagccacaagaagtag 282  
DB 112 GCAGCCCAAAAAAGCCATGAGAAAGCCACAAAGTAG 71

RESULT 14 389 bp mRNA linear EST 22-MAY-2000  
AW873740  
LOCUS hq29e11.x1 NCI-CGAP\_Adr1 Homo sapiens cDNA clone IMAGE:3120812 3'  
DEFINITION similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);, mRNA sequence.  
ACCESSION AW873740  
VERSION AW873740.1 GI:8007793  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS 1 (bases 1 to 389)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Chris Moskalko, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:  
info@image.lnl.gov  
Seq primer: -40UP from GIBCO.

FEATURES  
Location/Qualifiers  
1..389  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3120812"  
/clone\_lib="NCI-CGAP\_Adr1"  
/tissue\_type="neuroblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: adrenal gland; Vector: PCMV-SPORT6; Site\_1:

/note="Site\_2: Salt; Cloned unidirectionally. Primer:  
Oligo dT. Average insert size 1.2 kb. Library  
constructed by Life Technologies."

BASE COUNT 84 a 102 c 87 g 116 t  
ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 389;  
Best Local Similarity 100.0%; Pred. No. 1.8e-66;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgttgacgcagctgagaaagccttgaaactatcatcgacgcttaccacaagctacc 60  
DB 352 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTTATCATCGACGTACCAAGTACTCC 293

OY 61 ctgataaagggaattccatgcgctctacagagatgacctgaagaattgctagaacc 120  
DB 292 CTGATTAAGGGGAATTTCCATCGCTACAGGATGACCTGAAGAATTTGCTAGAGACC 233

OY 121 gagttctcagatatacaggaagaggtgcagacgtctggttcaaaagattgatac 180  
DB 232 GAGTGTCTCACTATATACGAGAAAGGTCGACACGTCTGTTCAAGAGTTGATATC 173

OY 181 aacactgattgtcagtttaacttccagagattccctcatcttggtatagaatggcggt 240  
DB 172 AACACTGATGTGTGCACTTACTTCAGAGATTCCTCATTTGTTGATTAAGTGGCGGTG 113

OY 241 gcagcccaaaaaaagccatgaagaagccacaagaagtag 282  
DB 112 GCAGCCCAAAAAAGCCATGAGAAAGCCACAAAGTAG 71

RESULT 15 390 bp mRNA linear EST 25-JUN-1996  
W79564  
LOCUS W79564  
DEFINITION W79564.1 r1 Soares\_fetal\_heart\_MNH19W Homo sapiens cDNA clone  
IMAGE:346628 5' similar to gb:A12027\_cds1 CALGRANULIN A (HUMAN);,  
mRNA sequence.  
ACCESSION W79564  
VERSION W79564.1 GI:1390835  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 390)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maitra, M., Parsons, J.,  
Rifkin, B., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston  
R., Williamson, A., Wohlmann, P. and Wilson, R.  
The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Seq primer: mbd.RGA+ET  
High quality sequence stop: 375.

FEATURES  
Location/Qualifiers  
1..390  
/organism="Homo sapiens"  
/db\_xref="GDB:1272003"  
/db\_xref="taxon:9606"  
/clone="IMAGE:346628"  
/clone\_lib="Soares\_fetal\_heart\_MNH19W"  
/sex="unknown"  
/dev\_stage="19 weeks"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: heart; Vector: pRT73D (Pharmacia) with a









BASE COUNT 90 a 59 c 72 g 61 t  
 ORIGIN  
 /protein\_id="CAC16554.1"  
 /db\_xref="GI:1128053"  
 /translation="MTLEKALNSIIDVHKYSLIKGNFHAAYRDDLKULETECCQ  
 YIRKGAQVWKEKDINDGAVNQEFLILVIRKGVAAHKSHESHEKE"

Query Match 100.0%; Score 282; DB 6; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 6e-69;  
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 atgttcacgagctggagaagccttgaaactatcatcagacgtctacacagaactacc 60  
 Db 1 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATGAGCTGTACACAAAGTACGCC 60  
 QY 61 ctgataaagggaattccatgcgctctcagaaggtacgtctgaagaattctcagaacc 120  
 Db 61 CTGATTAAGGGGAATTTCCATGCCCTACAGGATGAGCTGAAGAAATTCGTAGAGACC 120  
 QY 121 gagtgcctcagatatatcagaagaaggtgcagacgctctgttcaagaagtgtatc 180  
 Db 121 GAGTGCCTCAGATATATCATGAGAAAGGCTGCAAGACGTCTGTTCAAAAGATTTGATATC 180  
 QY 181 aacactgatgtgcagtaacttcacagaggttcctcatctgtgtataaagatggcgtg 240  
 Db 181 AACACTGATGTGTCAGTTAACTTCCAGAGTCTTCATCTGTGATTAAGATGGCGCTG 240  
 QY 241 gcaagcccaaaaaaacatgaagaagcccaagaagtag 282  
 Db 241 GCAGCCCAAAAAAGCCATGAGAAAGCCACAAAGAGTAG 282

RESULT 2  
 LOCUS A12022 284 bp DNA linear PAT 11-JAN-1994  
 DEFINITION Macrophage migration inhibition factor (MRP-8).  
 ACCESSION A12022  
 VERSION A12022.1 GI:491243  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM  
 REFERENCE 1 (bases 1 to 284)  
 AUTHORS Odink, K.G., Clerc, R., Cerletti, N., Brueggen, J., Tarcsay, L., Sorg, C.  
 and Wiesendanger, W.  
 TITLE Novel lymphokine related peptides  
 JOURNAL Patent: EP 0263072-A 6 06-APR-1988;  
 CIBA-GEIGY AG  
 FEATURES  
 source Location/Qualifiers  
 1..284  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"

misc\_difference 1  
 /note="n= a flanking DNA residue of 12 nucleotides or more  
 containing a promoter sequence".

gene 2..283  
 /gene="MRP-8"

CDS 2..283  
 /gene="MRP-8"  
 /codon\_start=1  
 /transl\_table=1  
 /protein\_id="CAA00998.1"  
 /db\_xref="GI:491244"  
 /translation="MTLEKALNSIIDVHKYSLIKGNFHAAYRDDLKULETECCQ  
 YIRKGAQVWKEKDINDGAVNQEFLILVIRKGVAAHKSHESHEKE"

misc\_difference 284  
 /note="n= a flanking DNA residue of one or more  
 nucleotides or is absent"  
 BASE COUNT 90 a 59 c 72 g 61 t 2 others  
 ORIGIN

Query Match 100.0%; Score 282; DB 6; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 6e-69;  
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 atgttcacgagctggagaagccttgaaactatcatcagacgtctacacagaactacc 60  
 Db 2 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATGAGCTGTACACAAAGTACTCC 61  
 QY 61 ctgataaagggaattccatgcgctctcagaaggtacgtctgaagaattctcagaacc 120  
 Db 62 CTGATTAAGGGGAATTTCCATGCCCTACAGGATGAGCTGAAGAAATTCGTAGAGACC 121  
 QY 121 gagtgcctcagatatatcagaagaaggtgcagacgctctgttcaagaagtgtatc 180  
 Db 122 GAGTGCCTCAGATATATCATGAGAAAGGCTGCAAGACGTCTGTTCAAAAGATTTGATATC 181  
 QY 181 aacactgatgtgcagtaacttcacagaggttcctcatctgtgtataaagatggcgtg 240  
 Db 182 AACACTGATGTGTCAGTTAACTTCCAGAGTCTTCATCTGTGATTAAGATGGCGCTG 241  
 QY 241 gcaagcccaaaaaaacatgaagaagcccaagaagtag 282  
 Db 242 GCAGCCCAAAAAAGCCATGAGAAAGCCACAAAGAGTAG 283

RESULT 3  
 LOCUS AR015927 408 bp DNA linear PAT 05-DEC-1998  
 DEFINITION Sequence 2 from patent US 5776348.  
 ACCESSION AR015927  
 VERSION AR015927.1 GI:3972204  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM  
 REFERENCE Unclassified.  
 1 (bases 1 to 408)  
 AUTHORS Selengut, J.D., Orme-Johnson, W.H., Dretler, S.P. and Asakura, H.  
 TITLE Mineral precipitation system and method for inhibiting mineral  
 precipitation formation  
 JOURNAL Patent: US 5776348-A 2 07-JUL-1998;  
 FEATURES  
 source Location/Qualifiers  
 1..408  
 /organism="unknown"

BASE COUNT 118 a 91 c 106 g 93 t  
 ORIGIN  
 Query Match 100.0%; Score 282; DB 6; Length 408;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-69;  
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttcacgagctggagaagccttgaaactatcatcagacgtctacacagaactacc 60  
 Db 57 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCTATCATGAGCTGTACACAAAGTACTCC 116  
 QY 61 ctgataaagggaattccatgcgctctcagaaggtacgtctgaagaattctcagaacc 120  
 Db 117 CTGATTAAGGGGAATTTCCATGCCCTGTACAGGATGAGCTGAAGAAATTTGATAGACC 176  
 QY 121 gagtgcctcagatatatcagaagaaggtgcagacgctctgttcaagaagtgtatc 180  
 Db 177 GAGTGCCTCAGATATATCATGAGAAAGGCTGCAAGACGTCTGTTCAAAAGTGTGATATC 236  
 QY 181 aacactgatgtgcagtaacttcacagaggttcctcatctgtgtataaagatggcgtg 240  
 Db 237 AACACTGATGTGTCAGTTAACTTCCAGAGTCTTCATCTGTGATTAAGATGGCGCTG 296  
 QY 241 gcaagcccaaaaaaacatgaagaagcccaagaagtag 282  
 Db 297 GCAGCCCAAAAAAGCCATGAGAAAGCCACAAAGAGTAG 338

RESULT 4

193571 193571 418 bp DNA linear PAT 01-DEC-1998  
LOCUS Sequence 15 from patent US 5731166.  
ACCESSION 193571  
VERSION 193571.1 GI:3938041  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT 128 a 91 c 106 g 93 t  
ORIGIN

Query Match 100.0%; Score 282; DB 6; Length 418;  
Best Local Similarity 100.0%; Pred. No. 6.3e-69;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgttgaccgagctggaagaagccttgaactctatcatcagcgtctaccacaagtactcc 60  
57 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCATCATCGAGCTGACACAAAGTACTCC 116  
DB 61 ctgataaagggaattccatcagcgtctcagaagatgacctgaagaattctgataagacc 120  
117 CTGATTAAGGGGAAATTTCCATGCCGTCTACAGGAGTACCTGAAGAAATGCTAAGAGACC 176  
QY 121 gagtgccctcagatatacagaagaagggtgcagaagctgtgttcaagaagttgataacc 180  
177 GAGTGTCTCTCAGTATATCAGGAAAAAGGCTGCAAGAGCTGTGTTCAAGAAGTTGGATATC 236  
DB 181 aacactgatgtgagtttaacttccagaagttctctcatctcgtgtgataaagatggcgctg 240  
237 AACACTGATGTGGTCACTTAACCTTCAGAGATTCCTCATCTGCTGATTAAGATGGCGCTG 296  
QY 241 gcaagcccaaaaaaagccatgaagaagaccacaagaagtag 282  
DB 297 GCAGCCACAAAAAAGCCATGAAGAACCCACAAAGAGTAG 338

RESULT 5  
HSMRP8 418 bp mRNA linear PRI 12-SEP-1993  
LOCUS  
DEFINITION Human mRNA for calcium-binding protein in macrophages (MRP-8)  
ACCESSION X06234 Y00286  
VERSION X06234.1 GI:34772  
KEYWORDS calcium binding protein; cystic fibrosis antigen.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT  
The sequence of MRP-8 is identical to the sequence for the cystic

fibrosis CCR1 antigen, with one exception (additional G at 292; Y00278). This difference can reflect a natural variation in alleles or could arise from a sequencing error. X06234 sequence was confirmed for three independent cDNAs and for a chromosomal MRP-8 gene  
(lagasse E. and R.G.C., Mol. Cell. Biol., submitted).  
FEATURES  
source  
Location/Qualifiers  
1. 418  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="peripheral blood mononuclear cells"  
CDS  
57..338  
/note="MRP-8 (AA 1-93)"  
/codon\_start=1  
/protein\_id="CAA29580.1"  
/db\_xref="GI:34773"  
/db\_xref="SWISS-PROT:P05109"  
/translation="MTLEKALNINSDYHYKSLIKGNFHAVYRDDLKILETCPO  
YIRKQADWFKRIDINTDGAVNDFEFLIVIKMGVAARKKSHESHK"

BASE COUNT 128 a 91 c 106 g 93 t  
ORIGIN  
Query Match 100.0%; Score 282; DB 9; Length 418;  
Best Local Similarity 100.0%; Pred. No. 6.3e-69;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgttgaccgagctggaagaagccttgaactctatcagcgtctaccacaagtactcc 60  
57 ATGTTGACCGAGCTGGAGAAAGCCTTGAACTCATCATCGAGCTGACACAAAGTACTCC 116  
DB 61 ctgataaagggaattccatcagcgtctcagaagatgacctgaagaattctgataagacc 120  
117 CTGATTAAGGGGAAATTTCCATGCCGTCTACAGGAGTACCTGAAGAAATGCTAAGAGACC 176  
QY 121 gagtgccctcagatatacagaagaagggtgcagaagctgtgttcaagaagttgataacc 180  
177 GAGTGTCTCTCAGTATATCAGGAAAAAGGCTGCAAGAGCTGTGTTCAAGAAGTTGGATATC 236  
DB 181 aacactgatgtgagtttaacttccagaagttctctcatctcgtgtgataaagatggcgctg 240  
237 AACACTGATGTGGTCACTTAACCTTCAGAGATTCCTCATCTGCTGATTAAGATGGCGCTG 296  
QY 241 gcaagcccaaaaaaagccatgaagaagaccacaagaagtag 282  
DB 297 GCAGCCACAAAAAAGCCATGAAGAACCCACAAAGAGTAG 338

RESULT 6  
BC005928 423 bp mRNA linear PRI 12-JUL-2001  
LOCUS  
DEFINITION Homo sapiens, S100 calcium-binding protein A8 (calgranulin A),  
clone MGC:14536 IMAGE:4246359, mRNA, complete cds.  
ACCESSION BC005928  
VERSION BC005928.1 GI:13543538  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT  
Contact: MGC help desk  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) mcdexapil.stanford.edu  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 21 Row: c Column: 4  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 34772.

## FEATURES

## source

Location/Qualifiers

1..423

/organism="Homo sapiens"

/db\_xref="locusid:6279"

/db\_xref="taxon:9606"

/clone="MGC:14536 IMAGE:4246359"

/tissue\_type="Skeletal Muscle"

/clone\_id="MH\_MGC\_81"

/lab\_host="DH10b"

/note="Vector: pDNR-LIB"

56..337

/product="S100 calcium-binding protein A8 (calgranulin A)"

/protein\_id="AAH05928.1"

/db\_xref="GI:13543539"

/translation="MLTELEKALNSIIDVYHKSLIKGNPHAVYRDILKLETECPQ YIRKKGADVWFREKIDINTDGAVNFQEFILIVIMGVAAHKSHSHKE"

135 a 91 c 106 g 91 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 282; DB 9; Length 423;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-69;

Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttgaccgagctggaagaagccttgaactctatcatcagcgtctcaccaagctacc 60

DB 56 ATGTTGACCGGAGTGGAGAAAGCCTGAACTCATCATGACGCTGACCAAGTACTC 115

QY 61 ctgataaagggaattccatccgctctacagagatgacttgaagaatctcagaagacc 120

DB 116 CTGATAAAGGGGAATTTCCATGCGTCTACAGGAGATGACTGAAGAAATTTGTAAGAC 175

QY 121 gagtgcctcagatatacaggaagaagggtcagacgctcgttccaaagagttgatatc 180

DB 176 GAGTGTCTCCTCAGTATATCAGAGAAAGGTCGACGCTGCTGTTCAAGAGTTGGATATC 235

QY 181 aacactgatggtcagtttaacttccagagagttcccatcttggtatgaagaatggcgctg 240

DB 236 AACACTGATGTTGAGTAACTTCCAGAGCTTCCTCATTCGTGATAAAGATGGCGCTG 295

QY 241 gcaagcccaaaaaaagccatgaagaagccacaaagagtag 282

DB 296 GCAGCCCACAAAAAAGCCATGAAGAAAGCCACAAGAGTAG 337

## RESULT 7

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## BASE COUNT

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## Score

## DB

## Length

JOURNAL Patent: EP 0263072-A 10 06-APR-1988;  
 CIBA-GEIGY AG  
 Location/Qualifiers  
 1..485  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"

gene  
 124..405  
 /gene="MRP-8"

CDS  
 124..405  
 /gene="MRP-8"

124..405  
 /gene="MRP-8"

/codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA01000.1"  
 /db\_xref="GI:491248"  
 /translation="MLTELEKALNSIIDVYHKSLIKGNPHAVYRDILKLETECPQ YIRKKGADVWFREKIDINTDGAVNFQEFILIVIMGVAAHKSHSHKE"

BASE COUNT 152 a 106 c 114 g 113 t

ORIGIN

Query Match 100.0%; Score 282; DB 6; Length 485;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-69;  
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttgaccgagctggaagaagccttgaactctatcatcagcgtctcaccaagctacc 60

DB 124 ATGTTACCGAGCTGGAGAAAGCTTGAACCTATCATCATGACGCTTACCAAGTACTC 183

QY 61 ctgataaagggaattccatccgctctacagagatgacttgaagaatctcagaagacc 120

DB 184 CTGATAAAGGGGAATTTCCATGCGTCTACAGGAGATGACTGACCAAGAAATTTGCTAGAGACC 243

QY 121 gagtgcctcagatatacaggaagaagggtcagacgctcgttccaaagagttgatatc 180

DB 244 GAGTGTCTCCTCAGTATATCAGAGAAAGGTCGACAGCTGCTGTTCAAGAGTTGGATATC 303

QY 181 aacactgatggtcagtttaacttccagagagttcccatcttggtatgaagaatggcgctg 240

DB 304 AACACTGATGTTGAGTAACTTCCAGAGAGTTCCATTCCTGTTGATAAAGATGGCGCTG 363

QY 241 gcaagcccaaaaaaagccatgaagaagccacaaagagtag 282

DB 364 GCAGCCCACAAAAAAGCCATGAAGAAAGCCACAAGAGTAG 405

RESULT 8

LOCUS AX188193 2329 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 3888 from Patent WO0142467.

ACCESSION AX188193

VERSION AX188193.1

KEYWORDS AX188193.1 GI:15139666

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE Schlögel, R., Deeds, J., Berger, A. and Zhao, X.

AUTHORS Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer

JOURNAL Patent: WO 0142467-A 3888 14-JUN-2001;

REFERENCE Millennium Predictive Medicine, Inc. (US)

FEATURES Location/Qualifiers

1..2329

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 715 a 398 c 509 g 707 t

ORIGIN

Query Match 98.7%; Score 278.4; DB 6; Length 2329;  
 Best Local Similarity 99.6%; Pred. No. 7.7e-68;  
 Matches 279; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 gttgaccgagctggagaaagccttgactatcatcgcgtctacacaaagtactcc 62  
|||||  
Db 51 GTTGACCGAGCTGGAGAAAGCCTTGAACTCATCATGAGCTGTACCAAGTACTCCT 110  
QY 63 gataaagggaatttcacgtccgtctacagggtgagccctggaagaatttgtagagccga 122  
|||||  
Db 111 GATTAAGGGGAATTTCCATCCGCTTACAGGATGACCTGAATAATTTCTAGAGACCGA 170  
QY 123 gttgctcctagatatacagagaaaggtgcaagcgtctgctcaaaagtgtgataca 182  
|||||  
Db 171 GTGCTCAGATATATCAGGAAAAAGGTGCAGAGCTGTGTTCAAGAGTTGATATCA 230  
QY 183 cactgagtggtgagctgaactcagagagttccctcattcgtggtgataaagtggtgagc 242  
|||||  
Db 231 CACTGATGtGtGAGTtTACTTCTCTGAGtTCTCATTTGtGtATTAAGtGtGCGtGtG 290  
QY 243 agccacacaaaaaagccatgaagaagccacaaagagtag 282  
|||||  
Db 291 AGCCACAAAAAAGCCATGAGAGAAGCCACAAAGAGTAG 330

RESULT 9  
HSCFANT 420 bp mRNA linear PRI 12-SEP-1993  
LOCUS HSCFANT  
DEFINITION Human mRNA for cystic fibrosis antigen (CFag).  
ACCESSION Y00278  
VERSION Y00278.1 GI:29887  
KEYWORDS calcium binding protein; cystic fibrosis antigen.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 420)  
Dorin,S.R.  
TITLE Direct Submission  
JOURNAL Submitted (27-APR-1987) S. R. Dorin, Clinical and Population  
MEDLINE Cytogenetics Unit, MRC, Western General Hospital, Crewe Road,  
AUTHORS Edinburgh, Scotland, Great Britain  
REFERENCE 2 (bases 1 to 420)  
Dorin,J.R., Novak,M., Hall,R.E., Brock,D.J., Secher,D.S. and van  
AUTHORS Heyningen,V.  
TITLE A clue to the basic defect in cystic fibrosis from cloning the CF  
JOURNAL Nature 326 (6113), 614-617 (1987)  
FEATURES  
source Location/Qualifiers  
1..420  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="1q12-1q22"  
/clone="CFAB-9"  
/clone\_1b="1ambda gtl0.chronic myeloid leukaemia cDNA"  
52..336  
/note="CFag (AA 1-94)"  
/codon\_start=1  
/protein\_id="CAA68390.1"  
/db\_xref="GI:29887"  
/db\_xref="SWISS-PROT:P05109"  
/translation="MTLEKALNSITIDYVHKYSIKGNFNAVYVDLKKLLETCRO  
YIKKKADYVFKELDINTGAVNFOEFLILVTKMWMOTPKAMKKATKSS"  
85..175  
/note="non-EF-hand calcium binding site"  
misc\_feature 199..289  
misc\_feature /note="EF-hand calcium binding site"  
BASE COUNT 138 a 90 c 101 g 90 t 1 others  
ORIGIN

Query Match 95.7%; Score 270; DB 9; Length 420;  
Best Local Similarity 99.6%; Pred. No. 1.5e-65;  
Matches 281; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 atgtgaccgagctggagaaagccttgactatcatcgcgtctacacaaagtactcc 60  
|||||  
Db 52 ATGTGACCGAGCTGGAGAAAGCCTTGAACTCATCATGAGCTGTACCAAGTACTC 111  
QY 61 ctgtataagggaatttcacgtccgtctacagggtgagccctggaagaatttgtagagcc 120  
|||||  
Db 112 CTGATTAAGGGGAATTTCCATCCGCTTACAGGATGACCTGAATAATTTCTAGAGAC 171  
QY 121 gattgctcctagatatacagagaaaggtgcaagcgtctggttcaaaagtgtgatac 180  
|||||  
Db 172 GAGTGTCTCAGTATATACAGAAAAAGGTGCAGAGCTGTGTTCAAGAGTTGATATC 231  
QY 181 aacctgattgctgaactcagagagttccctcattcgtggtgataaagtggtgagc 240  
|||||  
Db 232 AACACTGATGtGtGAGTtTACTTCTCTGAGtTCTCATTTGtGtATTAAGtGtGCGtGtG 290  
QY 241 gcaagccacacaaaaaagccatgaagaagccacaaagagtag 282  
|||||  
Db 291 GCAGCCACAAAAAAGCCATGAGAGAAGCCACAAAGAGTAG 332

RESULT 10  
AX184529 573 bp DNA linear PAT 06-AUG-2001  
LOCUS AX184529  
DEFINITION Sequence 224 from Patent WO0142467.  
ACCESSION AX184529  
VERSION AX184529.1 GI:15135885  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 573)  
Schlegel,R., Deeds,J., Berger,A. and Zhao,X.  
TITLE Genes, compositions, kits, and methods for identification,  
assessment, prevention, and therapy of cervical cancer  
JOURNAL Patent: WO 0142467-A 224 14-JUN-2001,  
FEATURES  
source Location/Qualifiers  
1..573  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 173 a 109 c 154 g 137 t  
ORIGIN

Query Match 81.7%; Score 230.4; DB 6; Length 573;  
Best Local Similarity 99.6%; Pred. No. 2.3e-54;  
Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 51 caagtactccctgataaagggaatttcacgtccgtctacagagatgacctgaagaatt 110  
|||||  
Db 10 CAGGTACTCCCTGATTAAGGGGAATTTCCATCCGCTTACAGGATGACCTGAATAAT 69  
QY 111 gctagagaccgagtgctctcagatatacaggaagaaggtgcaagcgtctggtcaaga 170  
|||||  
Db 70 GCTAGAGACCGAGAGTGCCTCAGTATATCAGAAAAAGGTGCAGAGCTGTGTTCAAGA 129  
QY 171 gttgataaacaactgattgctcagtttaacttcagagagttccctcattcgtgataa 230  
|||||  
Db 130 GTTGATATACACAGTATGTCAGTtTACTTCTCTGAGtTCTCATTTGtGtATTA 189  
QY 231 gatggcgtgagccacacaaaaaagccatgaagaagccacaaagagtag 282  
|||||  
Db 190 GATGCGGTGGCAGACCCACAAAAAAGCCATGAAAGAACCCACAAAGAGTAG 241

RESULT 11  
AX186496 497 bp DNA linear PAT 06-AUG-2001  
LOCUS AX186496  
DEFINITION Sequence 2191 from Patent WO0142467.  
ACCESSION AX186496  
VERSION AX186496.1 GI:15137931

KEYWORDS	human.
SOURCE	
ORGANISM	Homo sapiens
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 497)
TITLE	Schlegel, R., Deeds, J., Berger, A. and Zhao, X.
JOURNAL	Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer
FEATURES	Patent: WO 0142467-A 2191 14-UN-2001; Millennium Predictive Medicine, Inc. (US)
source	Location/Qualifiers
	1..497
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	150 a 98 c 118 g 131 t
ORIGIN	

Query Match	77.4%	Score 218.4	DB 6	Length 497
Best Local Similarity	99.1%	Pred. 5.4e-51		
Matches	230	Conservative	0	Mismatches 1; Indels 1; Gaps 1;
Qy	51	caagtactccctgataaaggggaattccatgcgctctcaaggagatgaacctgaagaat	110	
	11			
	11			
	11			
	11			
Db	36	CAGGTACTCTCCGTGATTAAGGGGAATTTCCATCCCGCTACAGGATGACCTGAAGAAAT	95	
Qy	111	gctagaaccggagtgtccctcagtatatacggaaaagggtgcaagcgtctggtttaaaga	170	
	111			
	111			
Db	96	GCATGAGACCGAGTGTCTTCACGATATACGAAAAAGGTGACAGCGTGGTTAAAGA	155	
Qy	171	gttgaatatcaacaactgagtgtgcagtttaacttccaggagtttcctatcttgytgataaa	230	
	171			
	171			
	171			
Db	156	GTTGGATATCAACACACGATGGTGCACTTAATCTCCAGAG-TCTCATCTCTGGTGATAA	214	
Qy	231	gatggcgtgtggcagcccaaaaaaagccatgaagaagccacaagaatag	282	
	231			
	231			
	231			
Db	215	GATGGCGTGGCGCCACAAAAAAGCCATGAAGAAGCCACCAAGAAGTAG	266	

RESULT	12
AX187395	
LOCUS	AX187395
DEFINITION	Sequence 3090 bp DNA linear PAT 06-AUG-2001
ACCESSION	AX187395
VERSION	AX187395.1 GI:15138853
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE	1 (bases 1 to 598)
AUTHORS	Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
TITLE	Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer Patent: WO 0142467-A 3090 14-JUN-2001;
JOURNAL	Millennium Predictive Medicine, Inc. (US) Location/Qualifiers
FEATURES	1..598
source	/organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	179 a 113 c 158 g 148 t
ORIGIN	

	Query Match	Similarity	Score	218.4	DB 61	Length	598
	Best Local	Similarity	99.1%	Pred	No. 5	5e-51	
	Matches	230	Conservative	0	Mismatches	1	Indels
							Gaps
							1-
Qy	51	caagtactccctgataaagggaatttcacatgcgcgtctacaggaatgacctgaagaatt	110				
	11						
	12	CAGGTATCCCTCG-TAAAGGGGAATTTCATGCCGCTACAGGGAATGACCTGAAGAATT	70				
Db	12	CAGGTATCCCTCG-TAAAGGGGAATTTCATGCCGCTACAGGGAATGACCTGAAGAATT	70				

QY	111	gctaaagacccagctgttcctcaaglatatacagaagaaaggttcagacgcctgtgttcaaga	170
Db	71	gctaaagacccagctgttcctcaaglatatacagaagaaaggttcagacgcctgtgttcaaga	130
QY	171	gttgatataccaactgtatgtgtgcaggttaacttccaagagttccctcatcttgtatataa	230
Db	131	gattgcatatcaaacactgaatgtgtgcaggttaacttccaagagttccctcatcttgtatataa	190
QY	231	gattgagcgtgtgtcagcccaacaanaaaagccatgaaagaaagccacaagaagttag	292
Db	191	gattgagcgtgtgtcagcccaacaanaaaagccatgaaagaaagccacaagaagttag	242

RESULT	13				
LOCUS	AX186192				
DEFINITION	AX186192	625 bp	DNA	linear	PAT 06-AUG-2001
ACCESSION	Sequence	1887	from Patent	WO0142467.	
VERSION	AX186192				
KEYWORDS	AX186192.1	GI:15137618			
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 625)				
JOURNAL	Schlegel, R., Deeds, J., Berger, A. and Zhao, X.				
FEATURES	Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer				
SOURCE	Patent: WO 0142467-A 1887 14-JUN-2001; Millennium Predictive Medicine, Inc. (US)				
LOCATION/Qualifiers	1..625				

	Query Match	77.4%	Score 218.4	DB 6:	Length 625;
	Best Local Similarity	99.1%	Pred No. 5-51:		
	Matches 230; Conservative	0;	Mismatches 1;	Indels 1;	Gaps 1;
<hr/>					
Qy	51 caagtaccctcgtataaaaggggaatttcacatgcgctccagagatgatcctaagaatt	110			
Db	39 CAGGTACTCCCTG-TAAGGGGAATTTCATGCCGCTCAGAGGAGACCTTAAGAATTT	97			
<hr/>					
Qy	111 gctagaaacccgagtgctcctaafatatacgaaaagggtgcagactctggtttcaaga	170			
Db	98 GCTAGAACCGAGTGCTCTCATATATCAGAAAAAGGTGCAGACGTGTGTTCAAGA	157			
<hr/>					
Qy	171 gttagacataacaactgatgtgtaacttccagaggtttccattctcgttgtataa	230			
Db	158 GTTGATATCAACACTGATGGTGCAGTTAACTCCAGGAGTTCTTATCTGTGTATAA	217			
<hr/>					
Qy	231 gatggcgctggcagcccacaaaaaaccaatgaagaagccacaagaatag	282			
Db	218 GATGGCGCTGCACCCACAAAACCCTTGAAGAAAGCCCAAAGAGTAG	269			

RESULT	14			
LOCUS	AX187557	381 bp	DNA	linear
DEFINITION	Sequence 3252 from Patent WO0112467.			
ACCESSION	AX187557			
VERSION	AX187557.1	GI:15139019		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 381)			
AUTHORS	Schlegel,R., Deeds,J., Berger,A. and Zhao,X.			
TITLE	Genes, compositions, kits, and methods for identification,			

JOURNAL assessment, prevention, and therapy of cervical cancer  
Patent: WO 0142467-A 3252 14-JUN-2001;  
Millennium Predictive Medicine, Inc. (US)

FEATURES  
Source Location/Qualifiers  
1..381

/organism="Homo sapiens"

BASE COUNT 116 a 78 c 90 g 97 t

ORIGIN

Query Match 70.6%; Score 199; DB 6; Length 381;

Best Local Similarity 96.6%; Pred. No. 1.5e-45;

Matches 225; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 51 caagctactccctgataaagggaattccatgcgtctacaggaatgacctaagaatt 110  
|||  
DB 10 CAGGTACTCCCTGATTAAGGGGAATTTCATGCCGTCTACAGGATGACCTGAAGAAATT 69  
QY 111 gctagaagaccgagtgctcctcagtatatcagaaagaagtgacagctctggttcaaga 170  
|||  
DB 70 GCTAAGAGACCGAGTGTCTCTCAGTATCAGGAAAAAGGAGCTGTGTTCAAGA 129  
QY 171 gttgatatcaacactgattggtcagtttaacttccaggaagttcctcattctgtgtataa 230  
|||  
DB 130 GTTGATATCAACACTGATGTCAGTTAACTTC--AGGAGTCTCATCTGTGTGATATA 187  
QY 231 gatggg-cgtggcagcccaaaaaagccatgaaagaagcacaagaagtag 282  
|||  
DB 188 GATGGCGGTGGCAGCCACAAAAAGCCATGAAGAAAGCCACAAAGAGTAG 240

RESULT 15

AX184691 AX184691 512 bp DNA linear PAT 06-AUG-2001

LOCUS Sequence 386 from Patent WO0142467.

DEFINITION AX184691

ACCESSION AX184691

VERSION AX184691.1 GI:15136053

KEYWORDS

SOURCE

ORGANISM

human.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..512

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 156 a 103 c 117 g 134 t 2 others

ORIGIN

Query Match 70.6%; Score 199; DB 6; Length 512;

Best Local Similarity 96.6%; Pred. No. 1.6e-45;

Matches 225; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 51 caagctactccctgataaagggaattccatgcgtctacaggaatgacctaagaatt 110  
|||  
DB 26 CAGGTACTCCCTGATTAAGGGGAATTTCATGCCGTCTACAGGATGACCTGAAGAAATT 85  
QY 111 gctagaagaccgagtgctcctcagtatatcagaaagaagtgacagctctggttcaaga 170  
|||  
DB 86 GCTAAGAGACCGAGTGTCTCTCAGTATCAGGAAAAAGGAGCTGTGTTCAAGA 145  
QY 171 gttgatatcaacactgattggtcagtttaacttccaggaagttcctcattctgtgtataa 230  
|||  
DB 146 GTTGATATCAACACTGATGTCAGTTAACTTC--AGGAGTCTCATCTGTGTGATATA 203

QY 231 gatggg-cgtggcagcccaaaaaagccatgaaagaagcacaagaagtag 282  
|||  
DB 204 GATGGCGGTGGCAGCCACAAAAAGCCATGAAGAAAGCCACAAAGAGTAG 256

Search completed: September 10, 2002, 06:42:40  
Job time: 6396 sec

7

---